Molecular ecological interactions of bacterial endophytes with their host *Vitis vinifera* (L)

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades

eines Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Juan Sebastian Lopez Fernandez

aus Bogotà / Kolumbien

1. Referentin: PD Dr. Barbara J. Schulz
2. Referent: Prof. Dr. Michael Steinert
eingereicht am: 22.03.2017
mündliche Prüfung (Disputation) am: 09.05.2017

Druckjahr 2017

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

1. Lòpez-Fernàndez, S. Sonego, P. Moretto, M. Pancher, M. Engelen, K. Pertot, I. Campisano, A. (2015). Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. Frontiers Microbiol. 6, 419.
2. Lòpez-Fernàndez, S. Compant, S. Vrhovsek, U. Bianchiedi, P L, Sessitsch, A. Pertot, I. Campisano, A. (2016). Grapevine colonization by endophytic bacteria shifts secondary metabolism and suggests activation of defense pathways. Plant Soil. 405, 155-175.
3. Lòpez-Fernàndez, S. Mazzoni, V. Pedrazzoli, F. Pertot, I. Campisano, A. (2017). A phloem-feeding insect transfers bacterial endophytic communities between grapevine plants. ***In review***, Frontiers in Microbiology.
4. Campisano, A. Pancher, M. Puopolo, G. Puddu, A. **Lòpez-Fernàndez, S.** Biagini, B. et al. (2015). Diversity in endophytic populations reveals functional and taxonomic diversity between wild and domesticated grapevines. Am J Enol Vitic. 66, 12-21.

Tagungsbeiträge

Lòpez-Fernàndez, S. Mazzoni, V. Pedrazzoli, F. Pertot, I. Campisano, A. “Adaptation of transmissible bacterial communities to multiple hosts: how the sap-feeding insect *Scaphoideus titanus* ships bacterial symbionts across grapevine plants”. XVII Congress Of The International Society For Molecular Plant-Microbe Interactions. OREGON CONVENTION CENTER. Portland – USA. (2016).
Lòpez-Fernàndez, S. Mazzoni, V. Pedrazzoli, F. Campisano, A. “Insect mediated transfer of microbial communities across plants”. COST action FA1405: “using three-way interactions between plants, microbes and arthropods to enhance crop protection and production”. ILUNION MALAGA HOTEL. Màlaga – Spain. (2016).
Lòpez-Fernàndez, S. Mazzoni, V. Pedrazzoli, F. Campisano, A. “Can insect pests be vectors of beneficial endophytes?”. International symposium “miCROPe” Microbe-assisted crop production- opportunities, challenges and needs. SCHLOSS SCHOENBRUNN TAGUNGSZENTRUM, APOTHEKERTRAKT. Vienna – Austria. November (2015).
Lòpez-Fernàndez, S. Compant, S. Vrhovsek, U. Bianchiedi, P L, Sessitsch, A. Pertot, I. Campisano, A. “Endophytic colonization of grapevine by bacteria reveals a metabolic signature suggesting activation of pathways for symbiosis and defense”. International congress “Rhizo4: stretching the interface of life”.WAGENINGEN UNIVERSITY: MAASTRICTH EXPOSITION AND CONFERENCE CENTER. Maastricht – The Netherlands. (2015)

Lòpez-Fernàndez, S. Sonego, P. Moretto, M. Pancher, M. Engelen, K. Pertot, I. Campisano, A. “A whole-genome comparison of virulence traits in endophytic genomes of Enterobacteria”. COST action meeting: Risk assessment of endophytes. UNIVERSITY OF EGE. Izmir - Turkey. (2014)

Posterbeiträge

Lòpez-Fernàndez, S. Campisano, A. Covelli, L. “Comparison Of Three Artificial Methods For The Re-Inoculation Of Bacterial Endophytes In Micropropagated *Malus Domestica* (Borkh) Plantlets”. XXII National Congress of the Italian Society of plant pathology. CENTRO DI RICERCA PER LA PATOLOGIA VEGETALE DI ROMA. Rome – Italy. (2016)

Lòpez-Fernàandez, S. Campisano, A. Schulz, B. “Fungal Endophytes From Grapevine Have Host-Dependent Levels Of Virulence And Produce Antibiotic Compounds In Dual Cultures”. XVII Congress of the international society for molecular plant-microbe interactions. OREGON CONVENTION CENTER. Portland – USA. (2016)

Lòpez-Fernàndez, S. Mazzoni, V. Pedrazzoli, F. Pertot, I. Campisano, A. “Adaptation of transmissible bacterial communities to multiple hosts: how the sap-feeding insect S*caphoideus titanus* ships bacterial symbionts across grapevine plants”. XVII Congress of the international society for molecular plant-microbe interactions. OREGON CONVENTION CENTER. Portland – USA. (2016)

Aknowledgements

I want to thank my family for an incredible long distance support during these 4.3 years away from home, for being my inspiration, motivation, reason, goal, point of reference, support system and more.

I’m truly grateful with Dr. Ilaria Pertot from Fondazione Edmund Mach for trusting me her facilities, materials and ideas to make the most out of my PhD research. Also, for spending a considerable amount of time to discuss, review, analyze and revisit the data obtained during my research. Also want to express my gratitude to Dr. Andrea Campisano who more than a supervisor was a mentor in microbiology. He designed most of the experiments included in this thesis. Also critically revised all the manuscripts and motivated me to analyze the data from multiple perspectives. He also was involved in financing my PhD through several collaborations worldwide and few words of gratitude would be unfair for such an incredible job as a supervisor.

Thanks to Dr. Barbara Schulz at the TU-Braunschweig for mentoring my way to understand plant-associated microorganisms and for sharing many important theoretical and practical secrets of endophytes. But most of all for believing in my capabilities. Also for being part of my *”Kommission”*.

I’m truly grateful for Prof. Dr. Marc Stadler at the Helmholtz Zentrum für Infektionsforschung letting me join the group as a guest student, and allowing me to have an incredible experience in biotechnology with fungi,

even when I am truly a bacteriologist. His excellence inspires me to follow his steps. Also for being part of my "*Kommission"*.
My most sincere gratitude to Dr. Frank Surup for being my mentor in secondary metabolites of fungi. Dr. Surup analized data from Mosher, feeding, NMR and HPLC–MS experiments. Especially want to thank Simone Heitkämper, Cäcilia Schwager, Axel Schulz, Silke Reinecke, Kerstin Schober and Mrs. Christiane Fritz-Braun (for always greeting me in Spanish!). Also a big thank you to Dr. Kathrin Mohr and her technicians for helping me in the bacteriology lab.

Thanks to Prof. Dr. Michael Steinert for letting me be part of the working group and use the facilities, equipment and expertise of the group to discuss my results. Also for being part of my *"Kommission"*. I don´t have enough words to thank the colleagues from the AG. Steinert. To (my very good friend) Hilger Jagau, Dr. Janine Rasch, Dr. Can Ünal, Katharina Kleilein, Hannes Beims, Quinxu Hu, Gabi Günther, and all the other Steinerts. Thank you for your so good support during my stay in Germany!

Also I'm truly grateful to Dr. Angela Sessitsch from the Austrian Institute of Technology for lending me her labs and supporting me in performing experiments on FISH and colonization. In the same institute, I'm deeply grateful to Dr. Stéphane Compant for his valuable lessons in microscopy and ecology of endophytic bacteria.

A big thank you to the members of the European Cooperation in Science and Technology (COST) action FA1103, particularly Dr. Caroline Schneider and to the staff at the International Research School in Applied Ecology (IRSAE) in Norway for all the financial support for my doctoral thesis. Their investment and contributions clearly made my training and experiments much more feasible and practical.

Also thanks to doctors Christian Cainelli, Gerardo Puopolo, Valerio Mazzoni, Maria Cristina Crava, Valerio Rossi, Kathrin Wittstein**,** and Soleiman Helaly for their advice, help and guidance in little everyday things. Thanks to Gabriella Tait, Michael Pancher, Pierluigi Bianchedi and so many others for advise, support, technical expertise and also for being terrific colleagues. Thanks to the management staff at FEM (Flavia Zanon, Alessandro Gretter, Elisabetta Perini, Gonzalo Cervantes) who always took care of my inability of understanding European regulations and for managing finances and bureaucracy for my doctorate, which I would have never be able to do on my own.

A heartfelt thank you to friends and colleagues to whom I probably don´t have enough words to express gratitude to. Dr. Laura Covelli for her advice and supervision in many molecular biology experiments, but most of all for her friendship. To Fede for his patience and love in this incredible journey in Europe. I owe you more than you could ever imagine. To my best buddy (Zeraye Mehari Haile), to my flatmates and good friends Muddu (Mridula Prasad) and Rimmie (Rupinder Kaur), to Sudarshan Pradmasu, Santosh Revadi, Lidia Nicola and Lorena Herrera for becoming good friends that helped me survive in the Italian mountains. Special thanks to Dr. Carol Ximena Lopez for incredibly helpful hand on R but overall for her friendship and the grotesque amounts of laughter that we have shared. To Zeljka Rupcic, Clara Chepkirui, Sandra Halecker

Lucile Wendt, Christian Richter and all the other MWIS members for being really nice people. To Doctor Carlitos, “huesitos”, Serginho, Johnson, Migue and especially to “chiqui” for all their love that I believe is true.

Finally, I want to mention my high school biology teacher Paola Carolina Pena. Once she thought it would be important for us to hear about how biology had changed with the recombinant DNA technologies (that was back in the year 2000) and how bacteria were important for the advance of technology. Although at that time, a lot was unknown to me, she ignited my curiosity to the very point of helping me choosing a career in microbiology. If you ever read this, this is a tribute to you and I dedicate this work to your efforts to make our country wiser, more educated, more curious and safer for people that sometimes like many of us, have few opportunities to succeed. The role of teachers at school is incredibly underrated, under-appreciated and underpaid in Colombia. We all hope for the day when this incredible job becomes a priority for governments in the developing world.

“*per aspera ad astra***”**

Lucius Annaeus Sèneca

TABLE OF CONTENTS

I. ABBREVIATIONS AND SYMBOLS 10

II. LIST OF FIGURES 12

II. SUMMARY 14

1. INTRODUCTION 15

1.1 Definition of endophyte and endophytism 16
1.1.1. Biology of grapevine endophytes 17
1.1.2. Taxonomy of endophytes from grapevine 18
1.2. General aspects of endophyte ecology 18
1.2.1. Bioprospecting with endophytes 19
1.2.2. Recognition and colonization in endophytic ecology 21
1.2.2.1. Tools for visualization of colonizing endophytes 22
1.2.2.2. Tools for detection of active interactions of endophytes with their hosts 22
1.2.3. Transmission of endophytes 23
1.2.3.1. Insects as endophyte vectors 24
1.2.3.2. Tools for studying the transmission of endophytic microbiota 25

2. OUTLINE AND AIMS 27

2.1. Outline 28
2.2. Aims 28

3. MATERIALS AND METHODS 29

3.1. Bacterial and fungal endophytic strains 30
3.1.1. Growth media and buffers 30
3.1.2. Sterilization method 30
3.1.3 Isolation methods 30
3.1.4. Organisms 30
3.1.4.1. Fungal and bacterial strains 31
3.1.4.2. Plants 32
3.2. Biological functions and bioprospecting of bacterial endophytes from grapevine 32
3.2.1. Experimental design 32
3.2.2. DNA barcoding by 16S rDNA sequencing 33
3.2.3. Bacterial automated ribosomal intergenic spacer analysis (B-ARISA) 33
3.2.4. Screening for enzymatic activity 33
3.2.5. Screening for quorum sensing activity 33
3.2.6. Screening of swimming & swarming 34
3.2.7. Screening for biocontrol activity 34
3.2.7.1. Dual-plate antagonism 34
3.2.7.2. Biocontrol of Botrytis cinerea 34
3.2.7.3. Biocontrol of Plasmapora viticola 34
3.2.8. Screening for plant growth promotion activity ex planta 35
3.2.8.1. 1-aminocyclopropane-1-carboxylate (ACC) - deaminase activity 35
3.2.8.2. Indole-3-acetic acid (IAA) production 35
3.2.8.3. Nitrogen fixation 35
3.2.8.4. Phosphate solubilisation 35
3.2.8.5. Siderophore production 35
3.2.8.6. Ammonia (NH_3) production 35
3.2.9. Screening for antibiotic resistance 36
3.2.10. Growth promotion testing in planta 36
3.2.11. Statistical analysis 36
3.3. Genome sequencing and analysis of bacterial endophytes 37
3.3.1. Experimental design 37
3.3.2. DNA extraction, genome sequencing and assembly 37
3.3.3. Annotation and subsystem analysis 37
3.3.4. Phylogeny 38
3.3.5. Whole-genome comparison 38
3.3.6. Clustering of orthologous families 38

3.3.7. Secretion system analysis *39*
3.3.8. CRISPRs and phage presence *39*
3.4. Transmission of bacterial endophytes using Scaphoideus titanus as vector ***39***
3.4.1. Experimental design *39*
3.4.2. DNA extraction, 16SrDNA amplification and pyrosequencing *41*
3.4.3. Bacterial 16SrDNA amplicon demultiplexing and statistical analysis *42*
3.4.4. Transmission and quantification of endophytes through qPCR *42*
3.5. Colonization assays using bacterial endophytes from grapevine ***44***
3.5.1. Experimental design *44*
3.5.2. Double labeling of oligonucleotide probes-Fluorescence in situ hybridization (DOPE-FISH) *44*
3.5.3. Metabolic profiling of colonized grapevine plants *45*
3.5.4. Statistical analysis *46*

4. RESULTS AND REMARKS **47**

4.1. Diversity in endophyte populations reveals functional and taxonomic diversity between wild and domesticated grapevines ***48***
4.1.1. Endophytic bacterial communities in wild and cultivated grapevine differ in diversity *48*
4.1.2. Endopytes are a rich source of enzymatics functions *50*
4.1.2.1. Growth promotion ex planta *50*
4.1.2.2. Tests for antibiotic resistance *50*
4.1.2.3. Enzyme production and competition for nutrients *50*
4.1.2.4. Biocontrol activity. *51*
4.1.3. Specific taxa differentiate wild and domesticated grapevine communities *52*
4.1.4. Remarks *52*
4.2. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria ***55***
4.2.1. Sequencing and assembly of endophytic genomes *55*
4.2.2. Comparison of genome structure in test and reference strains shows lifestyles and chromosome arrangement are linked *57*
4.2.3. Core and accessory genomes from endophytes are enriched in virulence factors *59*
4.2.3.1. Genus Enterobacter *59*
4.2.3.2. Genus Erwinia *61*
4.2.3.3. Genus Pantoea *61*
4.2.4. Unique gene functions in the different lifestyles *64*
4.2.5. Other virulence determinants: CRISPRs and phage sequences *70*
4.3. A phloem-feeding insect transfers bacterial endophytic communities between grapevine plants ***74***
4.3.1. Structure of the community in the tested grapevine holobiont *74*
4.3.2. Selected endophytes are transmitted between grapevine plants *75*
4.3.3. Insects change the community structure during passage from source to sink plants *76*
4.3.4. Endophytic community composition shifts in a host- specific manner *80*
4.3.5. Endophytes are acquired by insects through feeding and delivered to the stems of grapevine plants *83*
4.3.6. Remarks *83*
4.4. Grapevine colonization by endophytic bacteria shifts secondary metabolism and suggests activation of defense pathways ***87***
4.4.1. Colonization of bacteria visualized through DOPE-FISH *87*
4.4.2. Metabolome analysis reveals effect of endophytic colonization on plant's metabolism *89*
4.4.3. Remarks *92*

5. CONCLUSIONS AND FUTURE WORK **95**

6. BIBLIOGRAPHY **105**

I. ABBREVIATIONS AND SYMBOLS

ACC	1-aminocyclopropane-1-carboxylate deaminase
AHL	*N*-acyl homoserine lactone
ANOSIM	Analysis of similarity
ANOVA	Analysis of variance
AM	Arbuscular Mycorhiza
B-ARISA	Bacterial automated ribosomal intergenic spacer analysis
BRIG	Blast ring image generator
BSA	Bovine Serum Albumin
Ca.	*Candidatus*
Ca	Circa
CDS	coding sequence(s)
CFU	Colony forming units
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
CRT	CRISPR recognition tool
cv.	Cultivar
DOPE- FISH	Double labelling of oligonucleotide probes – fluorescent in situ hybridization
DGGE	Denaturing Gradient Gel Electrophoresis
DNA	DeoxyRibonucleic Acid
DSE	Dark Septate Endophyte
dpi	days post inoculation
eGFP	enhanced green fluorescent protein
EIC	Extracted Ion Chromatogram
fAME	fatty acid methyl ester chromatography
FDAA	1-fluoro-2,4-dinitrophenyl-5- L -alaninamide
FISH	Fluorescent *in situ* hybridization
GC content	Guanine-Cytosine content
GMO	Genetically Modified Organism(s)
hpi	hours post inoculation
HPLC	High Performance Liquid Chromatography
IAA	Indole acetic acid(s)
IGS	Intergenic sequence(s)
IPM	Integrated pest management
LC-MS	Liquid chromatography – mass spectrometry
LPS	Lipopolysaccharide
LPSN	List of Prokaryotic Names with Standing Nomenclature
MAMP	Microbe-Associated Molecular Patterns
MCA	Markov Cluster Algorithm
MHB	Mycorhization Helper Bacteria

MS-ESI	Mass spectrometry – electrospray ionization
MTI	MAMP-Triggered Immunity
NBS-LRR	Nucleotide-binding site – leucine rich repeat
NCBI	National Centre for Biotechnology Information
NH_3	Ion Ammonium
NGS	Next Generation Sequencing
OMP	Outer Membrane Protein
OTU	Operational Taxonomic Unit
PAMP	Pathogen-Associated Molecular Pattern
PBS	Phosphate Buffer Saline
PCA	Principal Components Analysis
PCR	Polymerase Chain Reaction
PMI	Plant-Microbe-Insect interaction
PROKKA	Prokaryotic Genome Annotation System
PRR	Pattern Recognition Receptors
pv.	Pathovar
QIIME	Quantitative insights into microbial ecology
QS	Quorum sensing
qPCR	Quantitative Polymerase Chain Reaction
RAST	Rapid Annotation Using Subsystem Technology
RNA	Ribonucleic Acid
SCFA	Short Chain Fatty Acids
Subsp.	Subspecies
TES	Tris EDTA sucrose solution
TOF	Time of Flight
T1SS	Type 1 secretion system
T2SS	Type 2 secretion system
T3SS	Type 3 secretion system
T4SS	Type 4 secretion system
T5SS	Type 5 secretion system
T6SS	Type 6 secretion system
UDG	Uracyl DNA Glycosidase
UPLC	Ultra performance liquid chromatography
UV	Ultraviolet wavelength range
16L: 8D h	Photoperiod of 16 hours light and 8 hours darkness
16SrDNA	Gene coding for the RNA of the ribosomal small subunit 16S.

II. LIST OF FIGURES

Table 1.	Bacterial strains used in this study
Table 2.	Characteristics of sequenced genomes from grapevine endophytes and reference genomes
Table 3.	Unique functions in the genomes of Enterobacteria
Table 4.	CRISPRs found in the genome of strain EnVs6
Table 5.	OTUs transmitted from source (SRC) to sink plants (STEMSNK and ROOTSNK) by *S. titanus*
Figure 1.	Map of sampling sites across Italy for endophyte isolation
Figure 2.	Experimental design of artificial transmission using S. titanus as vector
Figure 3.	Principal component analysis of biotechnological properties of endophytes from grapevine
Figure 4.	Biocontrol activity of bacterial endophytes from grapevine.
Figure 5.	PGP activity of bacterial endophytes from grapevine
Figure 6.	Whole-genome comparisons in three genera of Enteobacteria
Figure 7.	Core and accessory genome size in endophytic and non-endophytic bacteria
Figure 8.	Venn diagrams showing shared and unique genes in Enterobacteriaceae
Figure 9.	Relative abundance of OTUs assigned at the phylum level
Figure 10.	Alpha diversity metrics on transmission experiments
Figure 11.	Comparisons of statistically significant abundances per sample category in a rarified OTU table
Figure 12.	Principal coordinates analysis of a rarefied OTU Table at phylum level
Figure 13.	Quantification of eGFP gene copy numbers in the plants and insects used in transmission experiments
Figure 14.	Visualization of *Enterobacter ludwigii* with DOPE-FISH
Figure 15.	Visualization of *Pantoea vagans* with DOPE-FISH
Figure 16.	Visualization of *Sphingomonas phyllosphaerae* with DOPE-FISH
Figure 17.	Concentration of metabolites in control and endophyte-inoculated grapevine plants
Figure 18.	PCA of the first comparison between plants inoculated with strain EnVs6 and controls.
Figure 19.	PCA of the second comparison between plants inoculated with strain EnVs6 and controls
Figure 20.	Heatmap showing distribution and shift of concentrations of metabolites in roots and stems of *V. Vinifera* L
Figure 21.	Model of interactions in the grapevine holobiont

A number of appendix figures and tables can be found as an electronic attachment in a CD-ROM

.

Appendix table 1.	Growth media and buffers

Appendix table 2. Characteristics of sequenced genomes from endophytic isolates
Appendix table 3. Statistics assembly endophytic genomes sequenced
Appendix table 4. List of categories used s parameters for comparative genomic analysis
Appendix table 5. Presence/absence map genomes *Enterobacter*
Appendix table 6. Presence/absence map genomes *Erwinia*
Appendix table 7. Presence/absence map genomes *Pantoea*
Appendix table 8. Phages in genomes of endophytic bacteria
Appendix table 9. Secretion systems in genomes of endophytic bacteria
Appendix table 10. Abundances of the bacterial endophytic community of grapevine in transmission experiments
Appendix table 11. Diversity metrics statistical analysis in transmission experiments
Appendix Table 12. Kruskal-Wallis tests on differential abundance
Appendix Table 13. List of metabolites involved in colonization of endophytic bacteria of grapevine

Appendix figure 1. Summary screening tests bioprospecting with endophytic bacteria
Appendix figure 2. Number of plants and B-ARISA quantification
Appendix figure 3. Molecular phylogeny of strain belonging to the four most represented genera among cultivable endophytes
Appendix figure 4. Taxonomical relationships of endophytic test strains
Appendix figure 5. Synteny plots of endophytic test strains
Appendix figure 6. Rarefaction curves per sample category of OUT abundance in transmission experiment
Appendix figure 7. Ternary plots based on a rarified OUT table including STEMSNK
Appendix figure 8. Ternary plots based on a rarified OUT table including ROOTSNK
Appendix figure 9. Probe specificity for DOPE-FISH experiments

II. SUMMARY

Bacterial endophytes are organisms that live asymptomatically within plants. Some endophytic bacteria display plant protection properties, such as synthesis of plant hormones, production of antibiotics against pathogens and deterrents against herbivores. All these capabilities have drawn the attention of agricultural research and agricultural industry, because of their potential and consequently imminent use as biotechnologies for better crop management. Such capabilities are the result of complex ecological interactions that endophytes engage in, including those with other organisms in the plant ecosystem. Thus, understanding the interplay between bacterial endophytes and co-inhabiting organisms may provide means of better exploiting such beneficial attributes.

Functional analysis with enzymatic assays has been the classical method for studying the roles of endophytic bacteria in their plant hosts. Recently, sequencing of genomes and metagenomic studies have provided better insights into colonization, niche expansion and chromosome structure, all of which can be reflected in the beneficial properties of some endophytes. Additionally, metabolomics has allowed an in depth study of symbiosis, revealing the importance of chemical crosstalk with the plant.

Evidence shows that the microbiota can have a strong influence on host fitness. In plants, several studies propose the use of microbes as probiotics, since in some soils the existence of a defined, beneficial microbiota can prevent the development of plant disease (these are called “suppressive soils”). Since endophytes harbor many beneficial properties, questions arise on whether the use of the endophytic microbiota can have a positive impact on crops. Also, contrary to their pathogenic counterparts, little is known about the horizontal transmission of entire endophytic communities and their possible use as probiotics.

Thus, the aim of this thesis was to study the beneficial properties of endophytic bacteria isolated from grapevine and evaluate their genome structures to identify similarities and differences between these symbionts and plant pathogenic, free-living and epiphytic symbionts. Enzymatic and functional tests showed that a collection of more than 100 bacterial endophytes isolated from grapevine have beneficial properties for plants spanning from plant growth promotion to biocontrol against major grapevine pathogens. Using genome sequencing, chromosome organization was analyzed, revealing gene functions that might be important for plant – endophyte symbioses. Additionally, similarities between endophytic and non-endophytic bacteria at the genome level revealed the virulence potential that endophytes hold in their genomes, thus suggesting a role of ecological constraints in defining the lifestyle of these symbionts. Through 454 sequencing, the role of the insect, Scaphoideus titanus, as a vector was confirmed, showing that it could transfer entire endophytic communities between host plants and also indicating that both plant and insect hosts have an effect on bacterial endophytic community structure. Finally, with metabolomics, a possible “metabolic signature” exerted by endophytes on the grapevine plants they colonize was recognized and the ecological implications of this phenomenon on symbiosis was outlined.

In summary, endophytes from grapevine harbor a great biotechnological potential reflected in their enzymatic activities, their genomes and in their interactions with other members of the ecosystem. Thus, investigations should be continued to develop them into possible plant probiotics.

1. INTRODUCTION

Studies of molecular ecology in plant-associated microorganisms is flourishing, e.g. due to new molecular tools that provide novel insights into interactions between microorganisms. Such insights will enable the development of methods for meeting the increasing demand for new sources of plant protection strategies. The use of chemicals and heavy metals is a big concern for environmental safety (Wuana and Okieimen, 2011). Also, the uprising of pesticide resistance which results in ineffective treatment of plant diseases has pushed forward research in plant-microbe interactions (PMI). From a more philosophical point of view, research in this area comes from the urge to understand where pathogens come from, but also how we can mitigate their effects using naturally occurring antagonists (Wu et al., 2009). Plants are not isolated organisms. Rather, they associate with other organisms as symbionts, whereby the entire entity of plant and its microbiome is called the holobiont (Bordenstein and Theis, 2015). These microorganisms, i.e. bacteria, fungi, viruses, prions, protozoa and algae that are associated with plants have properties that are finely tuned to the plant's physiology. They partake in the plant's life cycle to such an important extent that without them the plant host could not survive (Turner et al., 2013). The community of microorganisms that play important roles in the plant, which is surveyed through 16SrDNA amplification and detection, is regarded as the microbiota of the plants. The collection of genes and genomes and episomes of the microbiota is known as the metagenome. When considering not only cells but their functions in the environment and the interactions with the surroundings, it is convenient to speak of the plant's microbiome (Whiteside et al., 2015).

1.1 Definition of endophyte and endophytism

The microbiota of plants includes those within-dwelling organisms that live in close association with their hosts. These microorganisms, termed endophytes, mostly thrive harmlessly inside the plant and in some cases their occupancy is accompanied by beneficial effects. The term endophyte from the Greek "endos" meaning inside and "phyton" meaning plant has been used for many years to define this internal, neutral-to-beneficial part of the plant's microbial community (Schulz and Boyle, 2006). An increase in resistance against plant pathogens, higher plant productivity and resistance to extreme conditions like salinity, soil acidity and dryness (Compant et al., 2005; Khan et al., 2012; Santoyo et al., 2016) among others, have been recorded as positive traits provided by these microbes. They may also benefit the host through hormone synthesis (Mercado-Blanco and Lugtenberg, 2014), which in turn regulates phenological phenomena like flowering and ripening.

Endophytism is a lifestyle that can be temporary or long-term, for example when the plants acquire endophytes straight from the seed (vertical transmission) and thus live their whole life cycle with the symbiont, while others can be infected only for a certain period of time (horizontal transmission) (Tintjer et al., 2008). In both cases, the endophyte attains its nutrients *in planta* from the plants. That is to say that those microorganisms that harmlessly spend at least part of their life cycle within the plant should be considered as endophytes (Hardoim et al., 2015). Functional assays have demonstrated the beneficial properties of endophytes, for example nitrogen fixation assays (Gupta et al., 2012), *in vitro* plant growth promotion (Oteino et al., 2015), antibiotic production (Bhore et al., 2013) among others Nonetheless, their ecology and interactions with the plant's microbiome are still largely undescribed.

1.1.1. Biology of grapevine endophytes

Grapevine is one of the most important crops in southern Europe, North and South America and a landmark crop of the agricultural industry. France, Italy, Spain and USA are the four largest producers of wine, accounting for 10.7% (USA) to 16.5% (France) of total world production in 2014 (The Wine institute reports http://www.wineinstitute.org/resources/statistics). The grapevine ecosystem is of great importance since it harbors a great biological diversity and recently has been recognized as provider of ecosystem services such as carbon sinks and landscape preservation (Brunori et al., 2016). It is no surprise that so many scientific conglomerates and industrial partners have targeted this crop as a goal for research and development of agriculture.

The natural history of endophytes in grapevine is not well known. Earlier studies on the origins of endophytes claimed that at the cellular level, pathogens and endophytes were different enough to be separated into two categories (Fisher and Petrini, 1992; Ryan et al., 2008). New research has shown that endophytic bacteria and fungi share virulence factors with plant pathogens (Adame-Álvarez et al., 2014; Junker et al., 2012). However, their mode of action can be completely contrary. A formal hypothesis of these observations states that endophytes might be hypovirulent pathogens that reside dormantly in the plant due to genetic constraints (Yu et al., 2015), and that under appropriate conditions they can reactivate virulence and become "pathogenic". In fact an extended version of this hypothesis proposes that endophytic fungi and bacteria live in a balanced antagonism where host defense and microbial virulence factors might be finely aligned with plant metabolism, as a survival strategy for both partners (Schulz et al., 1999; Schulz and Boyle, 2005). This "endophytic continuum" has been evaluated with endophytes being tested in host and non-host plants, showing that they can, under certain circumstances, express virulence factors (Junker et al., 2012).

The endophytic community of the grapevine has been selected through domestication. Campisano et al. (2014b) showed how a possible horizontal inter-kingdom transfer event of the human associated *Propionibacterium acnes* took place during the Neolithic during domestication of plants. The endophytic *P. acnes* type Zappae was found as a member of the endophytic community in the bark and pith cells of Italian vine plants.

It has been previously shown through community-ecology molecular tools such as DGGE and fatty acid methyl esterases (fAMEs) chromatographic profiles, that bacteria residing in the phyllosphere can enter the plant leaves and shoots, suggesting that a part of the endophytic bacterial community might come from the surface-adhering bacteria (West et al., 2010). In that respect, Compant et al. (2010) demonstrated how the plant growth-promoting rhizobacterium *Burkholderia phytofirmans* PsJN colonizes grapevine plants starting as a rhizoplane bacterium, but moves rapidly to the root cortex and the central cylinder of the roots. After six weeks post inoculation (w.p.i.), bacteria can be found in the xylem vessels.

Grapevine's endobiome largely consists of metabolically active symbionts and some of them hold plant protection properties as well as plant growth promotion capabilities (Mercado-Blanco and Lugtenberg, 2014; Samad et al., 2017). As an example, the endophytic *Pantoea* species whose role in biocontrol is nowadays well established (Smits, 2011), has also been found in grapevine plants (Bulgari et al., 2009). These active endophytes are affiliated with different taxa (Marasco et al., 2013). Among the biological activities that may be crucial for the host's well-being are included synthesis of Indole acetic acid (IAA), synthesis and release of siderophores, solubilization of phosphates through the production of acids, production of ammonia,

protease activity and exopolysaccharide production (Compant et al., 2005; Marasco et al., 2013; Pirttilä et al., 2004).

The endophytic community is also susceptible to exogenous forces that shape it and restructure it. A survey of the bacterial diversity of endophytes in grapevine suggests different cultivars, temperatures, edaphological properties, and even management strategies affect the abundances of particular taxa, and with that their functions in the plant host. Similarly, fungal endophytic communities of grapevine are also affected by the pest management strategy, where apparently the organic strategy favors persistence of fungal endophytes (Pancher et al., 2012). Pest management thus has a big influence on diversity of endophytic organisms even when coming from the same cultivars.

This has also been shown by finding a higher diversity of bacteria in the leaves of grapevine following treatment with antifungal agents, (Perazzolli et al., 2014). Location of the crops and seasonality may also play a role in the way the endophytic community is structured, since different sampling months and sites provide differently assembled endophytic communities (Ding and Melcher, 2016).

1.1.2. Taxonomy of endophytes from grapevine

A complete survey of the grapevine's bacterial microbiome from Portugal and Spain showed that the grapevine microbiome is dominated by Proteobacteria, followed by Firmicutes and Actinobacteria. The most abundant families were the Comamonadaceae, Enterobacteriaceae, Moraxellaceae, Neisseraceae Pseudomonadaceae, Sphingomonadaceae and Streptococcaceae (Pinto and Gomes, 2016; Zarraonaindia and Gilbert, 2015).

Other studies (Bulgari et al., 2009) agree with the above mentioned findings: Betaproteobacteria and Gammaproteobacteria are the most abundant taxonomic groups, where the genus *Comamonas* (Commamonadaceae) is relevant for the former class while some pseudomonads and a large number of enterobacteria, including the extensively studied biocontrol agent *Pantoea* sp., are more abundant in the latter. Also the phylum Firmicutes (where *Bacillus fastidiosus* and *B. inoslitus* are highlighted) and the Actinobacteria (with *Clavibacter* sp, *Curtobacterium* sp. and *Rhodococcus* sp. being key players) have been proposed as constituting the majority of the endophytic bacterial community.

Taxonomy of bacterial endophytes spans to a wide number of taxa, varying with the plant host. Microbial diversity of endophytes in rice plants from different cultivars worldwide is largely represented in the proteobacteria, with Alphaproteobacteria, Betaproteobacteria, Actinobacteria and pseudomonads (Hardoim et al., 2011). In noni (*Morinda citrifolia* L.), bacterial endophytes are included predominantly in the Proteobacteria, however other taxa, like Actinobacteria and Bacteroidetes, make a great proportion of the community (Yang et al., 2015). The same applies for the endophytic communities of wheat (*Triticum aestivum*) and corn (*Zea mays* L.) (Matsumura et al., 2015; Robinson et al., 2016).

1.2. General aspects of endophyte ecology

Endophytes can engage in lifestyles spanning from obligate plant symbionts, i.e. biotrophs, organisms that need plant tissues to survive, to facultative and opportunistic colonizers that are found mainly as epiphytes (Clay, 1993; Hardoim et al., 2015). In all three cases, endophytes may come from the microbiota residing in

the soil as free living microorganisms or as discussed above, from the phyllosphere. These can be recruited by plants, activated and assimilated as part of the host's microbiota (Bulgarelli et al., 2013). In such cases, plant hormone production, host genotype and soil composition strongly influence the numbers and taxa of endophytes to be selected (Gaiero et al., 2013).

Endophytes occupy a large array of niches (Matsumura et al., 2015). Bacterial and fungal endophytes each have their own defined functional groups. The endophytic community in rice plants, for example, contains bacteria with functions for nitrogen fixation, denitrification, nitrification and nitrate assimilation (Sessitsch et al., 2012). Other functions such as herbivore determent (Braun et al., 2003) have been documented. In another example, certain functional groups have been identified in endophytes from tomato where the community of root-associated endophytes contained enriched functions for carbohydrate and amino acid metabolism and was depleted in genes related to secondary metabolism, dormancy and sporulation (Tian et al., 2015). Nitrogen fixation and nitrogen metabolism were predominant in this example. Other examples show how some members of the endophytic community are responsible for coordination of metabolism through cell-to-cell signaling (Elasri et al., 2001). It is however important to highlight that endophytes ressemble other types of lifestyles in terms of gene functions. Capabilities of plant growth promotion might and biocontrol can also be present in free living organisms. But since endophytes are readily accessible in the plant, interest has grown to analyze them in small and big scale

1.2.1. Bioprospecting with endophytes

Endophytes as many other symbionts and non-symbionts can be seen as an immense source of gene functions accumulated through millions of years of co-evolution with plant hosts. A plethora of gene functions that are important for plant-microbe interactions has been described at the genomic level (Brader et al., 2014; Kaul et al., 2016). For example, secretion systems like T1SS and T2SS are abundant in genomes of endophytic bacteria, but T3SS and T4SS are rare. Also functions for 2,3-butanediol, acetoin and indole acetic acid production, which are important for the agricultural and biofuel industry, are abundant in endophytic bacterial genomes (Reinhold-Hurek and Hurek, 2011). Other functions protect the plant against variable environmental conditions. Among these functions are ice nucleation activity, the synthesis of 1-aminocyclopropane-1-carboxylate deaminase (ACC deaminase), trehalose synthesis (Chaturvedi and Singh, 2016) and molecular nitrogen fixation (Hurek and Reinhold-Hurek, 2003). Also, endophytes in grapevine possess a large number of genetic traits for iron uptake and chelation (TonB dependent receptors, hydroxymate synthases), plant cell wall-degrading enzymes and detoxification and chemotaxis (flagellum and T3SS, chemorreceptors) among others (Mitter et al., 2013). Although these genes may be present in the genomes of a wide variety of endophytes, activities within the same taxon but isolated from different hosts may also vary.

In bioprospecting, genomic studies have shed light onto organization and structure of bacterial chromosomes making possible, for example, the demonstration of operons and metabolic control systems (Binnewies et al., 2006) which translates into the development of new genetic engeneering tools for maximizing enzymatic production and high yield of metabolites. On the other hand, genomics permits tracing virulent and hypervirulent bacterial and fungal strains allowing the application of fast contingency plans for infection control.

Comparative genomics is a recent approach to study plant-microbe interactions and for bioprospecting. It involves the sequencing, annotation and parallel analysis of genomes from one or several organisms on the basis of a fully sequenced completely assembled and accurately annotated reference or template genome. An important feature of this approach is the ease with which information on the life history of microorganisms can be gained by comparing genomes of organisms diverging in lifestyle, habitat, niche and metabolism. For example, using this approach research has found that pathogens and obligate symbionts may occupy closely related niches, with sometimes overlapping functions (Suen et al., 2007). Also, genomic regions that are present or absent in two individuals of the same taxonomical rank inhabiting two different hosts could suggest adaptation events imprinted on the genome.

A practical goal of comparative genomics is to establish the core and accessory genomes, defined respectively as the **orthologues** (genes with a common ancester that have been derived through speciation) present in all the species of the genus under study, and the strain or species specific genes present in a given genome (Gabaldón and Koonin, 2013). Furthermore, comparative genomics intends to present the major rearrangements in the genome that might reflect lifestyle switching.

Comparative genomics of plant-associated bacteria shows that rearrangements such as symmetrical inversions and repetitive extragenic palindromic sequences (REP) populate their genomes, and multiple phages and mobile elements can also be found (Loper et al., 2012). A common characteristic in plant symbionts is the overrepresentation of amino acid and carbohydrate metabolism genes (Taghavi et al., 2010; Tian et al., 2012). Sequencing of genomes from plant-associated bacteria also show that core and accessory genomes may vary drastically in size, but usually span between 1500 to 3000 genes for the core genomes and a varied number of genes in the accessory genome.

Studies of comparative genomics of bacterial endophytes are still scarce and many questions regarding virulence potential, reorganization of the genomes during symbiosis and the transfer of genetic material from and to the host remain unanswered, although some efforts have been made to estimate organization of genomes, and to highlight genes that are important for survival inside the plant (Silby et al., 2009; Sugawara et al., 2013; Taghavi et al., 2009). In 2016, sequencing of the genome of a rice endophytic *Staphylococcus epidermidis* (RESE) showed a number of functional categories in its accessory genomes that seem to be important for association with the plant. Some of these are genes related to replication, recombination and DNA repair systems, as well as in functions for transport of inorganic compounds and ions (Chaudhry and Patil, 2016). Similarly, in 2010, Taghavi et al. showed that in the genome of *Enterobacter* sp. 638, an endophyte of poplar, about 19% of the coding sequences (CDS) were still lacking functional assignment. This shows how important genomic analysis is for the discovery of new metabolic pathways in endophytes and for the discovery of new biotechnological properties.

Genomics has also contributed to the understanding of the origin of endophytism. Sequencing of the genome of the rice fungal endophyte *Harpophora oryzae* showed that it is closely related to the rice pathogen *Magnaporthe oryzae*. The study proposed that the genome of the endophyte was larger than that of *M. oryzae* and that the rice endophyte shares a common ancestor with the pathogens *M. oryzae* and *M. grisea* along with *Nakataea oryzae*, but the lifestyle has been gained and lost randomly along the evolutionary track of the fungus (Xu et al., 2014).

1.2.2. Recognition and colonization in endophytic ecology

The interactions of endophytes and plants are a very fascinating example of co-evolution. Some estimates assume that endophytes and plants have been associated for more than 400 million years (Rodriguez and Redman, 2008). The colonization process of plants by endophytic fungi and bacteria provides clues as to how this bidirectional relationship has grown over millions of years.

Root colonization has been a focusing point in endophytic colonization research, since endophytes may come from the soil. The root acts as a chemical control tower that directs endophytes towards the plant. Root exudates of rice plants, for example, contain high concentrations of carbohydrates and amino acids that play a role as chemoattractants for endophytic bacteria (Bacilio-Jimenez, 2003). **Rhizodeposition**, the process of exudate production and micro-environment enrichment through root cap cell detachment is notably important in luring bacteria to the endosphere (Bulgarelli et al., 2013). In grapevine, the effect of root exudates on colonization may have an impact on the composition of root microbiota which apparently changes from bacteria- to fungi-dominated once a concentrated solution of exudate is applied *in vitro* (Li et al, 2013). On the contrary, only few studies have addressed colonization by bacteria of the aerial parts of the plant. For example, endophytes can infect through the leaves using stomata and fractures on the epidermis of the plant as entry points, although this seems less common (Compant et al., 2005).

At the molecular level, the cues for colonization have been revisited using the archetypal pathogenic colonization pathway as a template. In pathogens, recognition of plant involves immune responses dependent on Pathogen-Associated Molecular Patterns (PAMPs) that interact with Pattern Recognition Receptors (PRR) in the plasma membrane of plant cells that activates what is known as the MAMP-triggered immunity (MTI). This results in a reprogramming of gene expression during the interaction with the microbe (Jones and Dangl, 2006). Effectors injected inside the plant cells may interact with intracellular receptors called R proteins that will lead to hypersensitivity and localized cell death (Trdá et al., 2015).

Although the MTI response is also documented for endophytic colonization, the response diverges from that associated with pathogen colonization. Trdà and team (2013) have shown how flagellin (a MAMP) from *B. phytofirmans* is recognized by the grapevine receptor VvFLS2 (*Vitis vinifera* flagellin sensor), which allows the bacterium to harmlessly colonize the plant. This mechanism does not induce a proper immune response but rather a mild oxidative burst, a weak expression of defense genes (acidic chitinase, basic glucanase, protease inhibitor and lipooxygenase). There are no cellular changes associated with the pathogenic MITIresponse. Comparisons of gene expression profiles with a non-host *Pseudomonas syringae* pv *pisi* suggest that the defense gene expression is milder in plants inoculated with endophytic *B. phytofirmans* (Trdá et al., 2014).

Other mechanisms of colonization utilized by endophytes include **quorum sensing** (QS) and cell-to-cell recognition (Hartmann et al., 2014). These sensing mechanisms allow bacteria to synchronize gene expression and regulate exoproduct synthesis in most of the cases via *N*-acyl homoserine lactones (AHLs). In the endophytic bacterium *Serratia plymuthica*, QS regulates the expression of chitinases and proteases, enzymes required for infecting the host, and the synthesis of IAA, which regulates the host's metabolism. AHLs may also play a role in recognition of the plant by endophytes, since it is known that some plants are able to synthesize AHL analogues. For example, rosmarinic acid acts as a ligand for the RhlR receptor in the plant associated *Pseudomonas aeruginosa* PAO1 (Corral-Lugo et al., 2016).

1.2.2.1. Tools for visualization of colonizing endophytes

Nowadays, it is possible to artificially inoculate endophytes and track their colonization within the plant and thus recognize colonization patterns and tropisms (Bloemberg et al., 2000; Compant et al., 2010).

Basic uses of microscopy such as light microscopy with special dyes and more advanced techniques like transmission electron microscopy have been used to identify areas where colonization is higher (Torres et al., 2012)

Other techniques take advantage of DNA hybridization and polymerase chain reaction (PCR) to analyze endophytic cells in plants. For example, **fluorescent *in situ* hybridization** (Bulgari et al., 2011; Schmidt and Eickhorst, 2014) better known as FISH, has been used to discover the colonization pathway of endophytes. FISH is based on nucleic acid interactions with specific DNA or RNA probes. An advantage of this technique is the instantaneous visualization of microbes, but also the possibility to follow individual endophytic taxa at every stage. Also, when directed to mRNA, FISH permits the detection of metabolically active endophytes. Other more sophisticated tools have been used such as multi probe platforms known as “chips”, where several hundreds of probes of known genes from microbes are attached to a platform and thus a DNA extract of a community sample can be hybridized to the chip to define structure of the community. The “PhyloChip” (a platform where 16SrDNA probes of a large number of species are attached) for example has been used to study shifts in bacterial communities in response to infection of citrus plants with *Candidatus* Liberibacter asiaticus, which is a major plant pathogen in Asia (Sagaram et al., 2009).

1.2.2.2. Tools for detection of active interactions of endophytes with their hosts

With many plant associated bacteria, colonization starts by active movement, recognition, attachment and proliferation on or inside the plant.

Movement and **recognition** are two related processes, since bacteria possess receptors that will induce movement of flagella towards or away from the signaling molecule. *Agrobacterium* spp., for example, can recognize wounded plants, because they release chemicals (acetosyringone and hydroxysyringone) into the milieu that act as chemoattractants to the bacterium. These chemicals which are potent effectors (Mo et al., 1991; Gelvin, 2003) have been identified using mass spectrometry from extracts of plants. In rhizobia, association with the plant relies on the induction of *nod* genes, which are necessary for synthesis of an extracellular factor that controls reorganization of cells in the host (Long, 2001). In the symbiosis with *Sinorhizobium meliloti*, silencing studies using interference RNAs coupled to espectrometry techniques have shown the importance of flavones and flavonols of *Medicago truncatula* in the recognition and nodule formation by the bacterium (Zhang et al., 2009). Remarkably, luteoline, flavonoid that is easily isolated from plant cells, has been shown to play a key role in the correct functioning of the *nod* enzymatic apparatus for a proper interaction with the plant (Yeh et al., 2002). Another example of the role of plant-produced molecules in bacterial colonization can be found in the actinorhiza, where species of *Frankia* colonize the roots of plants in a process favored by auxin accumulation (Perrine-Walker et al., 2007). The mycorrhizal fungus *Glomus intraradices* can also recognize a group of well characterized molecules from the plant, known as strigolactones that have been identified using HPLC-MS (Oldroyd, 2013). *Pseudomonas syringae* recognizes the plant phenolic compound arbutin, which is necessary for colonization (Mo and Gross, 1991). In endophytes, chemical attraction and recognition is not completely understood. Shidore et al (2012) have

studied the responses of the rice endophyte *Azoarcus* sp. BH72 to the root exudates. Using a two-colored microarray, they showed how 176 genes are affected by the presence of the exudates, with most of the genes being upregulated. Among the gene functions that were affected by the root exoproducts they list genes for amino acid metabolism, energy production, cell wall and membrane biogenesis, and signal transduction mechanisms. Balachandar et al (2006) have also shown that plant flavonoids may affect colonization of plants by *Serratia* spp. In their study, the application of several flavonoids showed that the bacterium achieves efficient colonization when in contact with these molecules.

Attachment and **proliferation** of endophytes in the plant depends on the presence of pili, the pH of the microhabitat and the production of enzymes able to degrade tissues for deeper colonization. It has been shown that attachment and inner colonization depend on a molecular interaction between receptors on the plant and ligands on the microbe. Such evidence has been collected from experiments using mutants impaired in enzyme production or pili synthesis as well as mutants from host plants unable to produce the alleged receptors. The root hairs for example contain lectins that can be bound to bacterial outer membrane proteins (OMP) and thus facilitate attachment. Also, some bacteria are able to produce calcium dependent adhesins that recognize glycoproteins anchored to plant membranes and which initiate the deeper colonization process (Rodríguez-Navarro et al., 2007).

Compant et al (2008) used the endophytic bacterium *Burkholderia phytofirmas* PsJN as a model organism for studying colonization. Their study using engineered bacteria with inducible *lac* promoters, which suggests that the endophytic bacterium initially colonizes the root surface of grapevine plantlets just 3h p.i. and that the population (in terms of cell numbers) remains constant within the plant for five weeks. The bacterium can eventually be found in the root hair zone, in the root tips and in the branching zones of the secondary roots. The same group (Compant et al., 2005) has shown that the rhizoplane can contain an endophytic population of *B. phytofirmans* with as many as 4log CFU/g fresh weight, though as expected, the number decreases to less than 1log CFU/g fresh weight after the bacterium reaches the internal tissues of roots.

1.2.3. Transmission of endophytes

Endophytic microorganisms don't live isolated in the agroecosystem. Instead, constant exchange of matter and energy keep a balance in the system. To achieve colonization, endophytes have to reach plants aided by mechanisms including **pneumocoria** (transport by wind for sporulating endophytes), **zoocoria** (vectoring by animals and in the case of endophytes particularly through insects) and **water transport** in the form of aerosols. Other mechanisms like **seed infestation** and natural transport through seed feeding animals, may also contribute to endophyte dispersion (Truyens et al., 2015). For fungal endophytes, transmission can happen when the hyphae grow into the roots or the above-ground tissues of the plant, sometimes reaching the inflorescence and be transmitted vertically or horizontally when sexual or asexual spores are removed from the plant surface (Saikkonen, 2004)

In bacterial endophytes the major known transmission source comes from studies of seeds, for which it has been proposed that a large number of plants utilize **seed symbiosis** as an adapting mechanism. A long list of plant hosts bearing seed endophytes includes ash species (*Fraxinus* spp.), bean (*Phaseolus vulgaris*), coffee (*Coffea arabica*), grapevine (*Vitis vinifera*), maize (*Zea mays*), pumpkin (*Cucurbita pepo*), rice (*Oryza sativa*), the norway spruce (*Picea abis*), tobacco (*Nicotiana tabacum*), and a large number of grasses such

as *Lollium multiflorum* and *Panicum virgatum* (Truyens et al., 2015). Espinosa-Urgel et al (2002) have shown that attachment of *Pseudomonas* to corn seeds depends on calcium binding proteins and some virulence factors (toxins and hemolysins). Compant et al (2011) also showed how in inflorescences as well as in seeds from grapevine, bacterial endophytes are at large, reaching population densities of 3.43 log10 CFU/g and 1.44 log10 CFU/g respectively as compared to the populations in the rhizosphere which could reach up to 6.89 log10 CFU/g.

Besides plants, endophytes have close contact with animals, especially insects, which are one of the major plant visitors (Raman et al., 2012). Insects interacting with endophytic fungi, for example, may benefit by acquisition of nutrients such as sugars and nectars produced by the fungus, while the endophyte can horizontally reach hosts other than their primary host plant (Barelli et al., 2016). In this case, colonization by insect-transmitted endophytes will also benefit the plant when the fungus displays PGP or produces metabolites that are toxic to phytopathogens.

On the other hand, some endophytes can synthesize secondary metabolites that are deterrent to herbivorous insects (Raman et al., 2012; Pažoutová et al., 2013). In such cases, if the plant is protected from the herbivore by the fungus, the fungus can complete its life cycle inside the host, while the insect is being controlled by metabolites the endophyte secretes in the holobiont (Clay, 1993).

Although readily available information about fungal endophytes and insect-plant interactions populate the databases, to our surprise very few examples of bacterial endophytes in plant-microbe-insect interactions (PMI) have been reported. Thus, our studies are a well expected contribution to the subject where information concerning prokaryotes is still missing.

As pointed by Bennett (2013), the study of plant-insect-microbe interactions, in particular that of endophytic bacteria is of utmost importance since simultaneous synergistic, antagonistic and additive interactions are occurring that are determintal for plant protection and correct agricultural development. An example cited by the author is the transmissibility of plant pathogens and plant beneficial microbes through insects. The understanding of the routes of transmission can lead to effective means of controlling insect populations that act as vectors of pathogenic microbes and thus naturally control invasive microbial communities that might potentially be harmful to crops.

1.2.3.1. Insects as endophyte vectors

In grapevine crops, a large number of insects may play both beneficial and detrimental roles. Some of the most important insects that are associated with grapevine in Europe are the European grapevine moth *Lobesia botrana* (Denis & Shiffermüller 1775), the cicad of boisnoir *Hyalesthes obsoletus* (Signoret 1865) and the grapevine beetle *Pelidnota punctata* (Linnaeus 1758). These insects are recognized as pests for the crops, since they feed on the leaves and have several ways in which they interfere with grapevine reproductive stages (Reineke and Thiéry, 2016). These visitors occasionally accompany grapevine during its development, and have a rather wide host range.

There are, however, other insects whose development or sexual cycles depend on the host plant and have a greater impact, since they are well known carriers of plant pathogens and endosymbionts. This is the case with the American grapevine leafhopper *Scaphoideus titanus* (Ball, 1932), the tea green leafhopper

Empoasca vitis (Goethe, 1875) and the burning grape leafhopper *Jacobiasca lybica* (Bergevin & Zanon 1922), which are all known as leafhoppers and belong to the family Cicadellidae.

Scaphoideus titanus is a recently introduced vine pest in Europe. First found in France in 1958, this leafhopper imported from the United States is becoming a major concern for vineyards in southern Europe as it still is in North America. Its association with wine plants starts with egg laying in the excoriated bark of wood vines where the larvae will hatch ca. six months after the laying period (Chuche and Thiéry, 2014). Eventually, eggs will go through a diapause stage and hatch on the bark, where larvae are liberated and start growing through five developmental stages until adulthood. Then, they reproduce and stay alive for one further month.

The importance of *S. titanus* in the ecology and management of vineyards results from its ability to very efficiently vector diseases of grapevine. Larvae will feed on the sap of grapevine during the early stages after hatching. Feeding occurs by probing phloem from leaves after perforation between two cells is made by the sheath covered stylets. During this process, larvae can acquire pathogens like *Ca.* Phytoplasma vitis, a major vector-borne disease in Europe and North America (Chuche and Thiéry, 2014). Studies on the transmission of *Ca.* Phytoplasma vitis show that the bacterium can only be transmitted by feeding of the insect in the plant and is not transmitted to the eggs. In fact, studies show that in order to become an active carrier of the pathogen, the insect must constantly feed on infected plants in the first developmental stages, as the phytoplasma can be readily controlled in the insect's gut. After successful establishment in the insect's haemolymph, the phytoplasma can be injected into healthy plants in a passive form.

In the case of *E. vitis,* association occurs after overwintering on surrounding evergreen plants and migrating to the vineyard in early spring, where the insects develop for approximately four sexual generations (Decante and van Helden, 2006). *E. vitis* feeds from grapevine by puncturing the phloem vessels, enhancing an autonomous defense response in the plant, leading to necrosis and delayed maturity. Other effects of *E. vitis* in vine plants are a reduction of both photosynthetic rate and concentration of sugar in grapes (Fornasiero et al., 2016). It is well known that *E. vitis* can transfer bacterial symbionts of the genus *Cardinium*. Similar to what happens with *S. titanus*, after feeding from the phloem of the plant, *E. vitis* also transfers bacteria (Gonella et al., 2015). *Cardinium* is well known for being able to manipulate the reproductive system of insects. However, little information regarding its safety for grapevine plants is available.

There is also little information available regarding the transmission of bacteria to plants by *Jacobiasca lybica*. Nor has the life cycle of this insect been as extensively studied as the above mentioned. However, considering the feeding habits of this insect, transmission could be a possibility.

1.2.3.2. Tools for studying the transmission of endophytic microbiota

One major consideration of good agricultural practice is the increasing demand for environmentally friendly plant protection strategies, including a reduction in energy costs for the production of agrochemicals. Moreover, agricultural biotechnology is focusing nowadays on the development of prophylactic tools against bacterial, fungal and viral infections, just as medical biotechnology has been (Fox, 2015). For example, genetically modified (GM) plants are available that synthesize toxins against major herbivorous insects. This is the case with the so called Bt plants that encode recombinant versions of the *Bacillus thuringiensis* Cry proteins that target receptors in the midgut of *Ostrinia nubilalis,* the European corn borer (Betz et al., 2000).

Some other genetically modified plants are able to synthesize specific enzymes that reduce damage by pathogens. Such is the case with the wheat variety Bobwhite that harbors the *PmB3* defense transgene [a member of the coiled-coil nucleotide binding site leucine-rich repeat (NBS-LRR) type of disease resistance genes] conferring resistance against the powdery mildew *Blumeria graminis* (Zeller et al., 2013). The major concern with GMOs is the lateral transmission of transgenes to non-target species, the activation of signaling pathways that can be harmful for final consumers of vegetable products, the acquisition of resistance by pathogens and the ecological costs that these modifications may carry. Ecosystems may become destabilized by the introduced pressures to the already present pray-predator systems (Zeller et al, 2013). Harnessing the microbiomes of plants is a novel approach to stimulate plant defenses and could become a way of providing positive traits to plants (resistance to pathogens, bigger and more productive varieties, resistance to drought among others) without using genetic engineering tools (Mueller and Sachs, 2015).

Recently, the use of high throughput sequencing efforts like **next generation sequencing** (NGS) platforms have reduced costs and improved the capacity and speed of sequencing in the context of large numbers of samples. In endophyte research, many studies have used NGS to show the community structure in rhizobacteria and of endophytes in economically relevant plants (Akinsanya et al., 2015; Carrell and Frank, 2015). The importance of using NGS for community analyses relies also on the development of tools for microbial community analysis such as processors, software and on line assistance, like Mothur, QIIME, SILVA, etc.

One question that arises when thinking about microbiome manipulation is how a selected beneficial community can be delivered to plants. Several options have been considered in the past, using formulations of microorganisms and their spores based on biodegradable substances such as alginates and xanthan gums. A disadvantage of using cultivable microorganisms is that they could lose their functions when separated from the rest of the community. A very interesting strategy only recently proposed is the use of natural carriers to deliver symbionts. And insects are the first candidates to appear on the list of possible vectors.

2. OUTLINE AND AIMS

2.1. Outline

This research thesis was developed in the framework of the PAT (Provincia Autonoma di Trento, Italy) funded project “Endophytes: a new therapy for healthy, high quality grape” and a collaboration with the Helmholtz Zentrum für Infektionsforschung (Germany) in the department “Microbial Drugs”.
Since endophytes possess so many traits that could be exploited in agricultural biotechnology, we started investigations based on the premise that these symbionts might be a good treatment for plant diseases without involving genetically modified organisms and by taking advantage of naturally occurring processes for transmission, colonization and establishment. In particular, we wanted to take a closer look at the endophytic bacterial communities of grapevine, given that for some years now, scientists have focused on the manipulation of microbiomes to mitigate diseases and in the case of crops, increase productivity using methods that are less toxic and less invasive than synthetic chemical fertilizers are. Although harnessing the microbiome of plants is already a common goal in different research groups, we wanted to test the possibility to naturally deliver endophytes in a manner that would not create imbalances in the vineyard agroecosystem. In that sense, we hypothesized that to achieve a great reach in the fields, we should utilize naturally present vectors of endophytes that could spread the microorganisms across plants and then assess what the consequences of colonization *in planta* are once the endophytes reach their target host. However, a major concern nowadays is the biosafety of biofertilizers and microbe-based products. Thus, it was important to also dive deep into the genomic characteristics of endophytes to see to what extent further evidence could be provided for endophytism as a beneficial trait of bacteria.

2.2. Aims

With these premises in mind, our aims were:
1. To tests biotechnological properties of a collection of endophytic bacteria isolated from Italian grapevine, select some and sequence and analyze their genomes in order to understand symbiosis determinants and to evaluate biosafety issues (virulence traits) for potential application in the field.

2. To evaluate the transmissibility of endophytic bacterial communities of grapevine through the American leafhopper *Scaphoideus titanus* Ball. in an artificial experimental setting.

3. To assess the colonization of microhabitats by selected bacterial endophytes and decipher the effect of such colonization on grapevine secondary metabolism.

3. MATERIALS AND METHODS

3.1. Bacterial and fungal endophytic strains

3.1.1. Growth media and buffers

Several media were used in these studies. A comprehensive (but not exhaustive) list of media and buffers are listed in Appendix Table 1. Buffers from PCR experiments are not included, since they are part of kits that will be cited following instructions of the manufacturer.

3.1.2. Sterilization method

Where stated, plant material (stems leaves and roots of micropropagated plants) was surface-sterilized by successive washing in 98% ethanol for two min, 4% sodium hypochlorite for 2 min and 70% ethanol for 2 min as described previously (Pancher et al., 2012), and then rinsed three times with distilled sterile water. The water from the final washing step of all samples was plated on Luria Bertani agar (LBA) and incubated for 5 days at 30°C to check for microbial growth as a proxy for surface disinfection efficacy.

For biocontrol experiments, berries from grapevine plants were surface sterilized by soaking for 5 min in a 0.5 % NaOH solution, then washed three times with distilled sterile water and left to dry. The water from the final washing step of all samples was plated on LBA and incubated for 5 days at 30°C to check for microbial growth.

3.1.3 Isolation methods

For isolation of bacteria, approximately 1.5 m long vine shoots were surface sterilized as described above (see 3.1.2.). The bark was removed using a sterile scalpel blade under laminar air flow, and a 2 g section was coarsely ground in 1 ml of sterile saline solution (SSS). After grinding, samples were vortexed for 15 min, and 100 µl of the suspension were spread on Nutrient Agar (NA) medium (Oxoid, United Kingdom) amended with 5 mg/L of cycloheximide (Oxoid, United Kingdom) to prevent growth of fungal endophytes. Plates were incubated at 25°C for 48 h. Colonies appearing on the plates were further streaked on nutrient agar (NA). The purified isolates were grown overnight on nutrient broth (NB) medium including 5% glycerol and stored at -80°C for later use. At the time of experiments all the strains were recovered on NA and maintained fresh by transferring on new medium every 5 days.

Isolation of fungal endophytes followed the same procedure, but the isolation medium (PDA) was amended with chloramphenicol.

3.1.4. Organisms

All strains used in this study are listed in Table 1. Endophytes were isolated in the sampling points shown in Figure 1. Only the strains that were selected for deep analysis (i.e. genome sequencing, transmission and colonization experiments) are listed. The other isolates are briefly mentioned in Appendix Figure 1. All bacterial and fungal strains belong to Fondazione Edmund Mach and/or to the Helmholtz Zentrum für Infektionsforschung.

3.1.4.1. Fungal and bacterial strains

Table 1. Bacterial and fungal strains used in this study

STRAIN	ORIGIN	LIFESTYLE	USE
Achromobacter sp. AcVs1	*V. vinifera silvestris*	Endophytic	Genome sequencing, dual cultures
Enterobacter ludwigii EnVs2	*V. vinifera silvestris*	Endophytic	Genome sequencing
En. ludwigii EnVs6	*Vitis vinifera silvestris*	Endophytic	Genome sequencing, dual cultures, qPCr experiments, Colonization assays, Targeted metabolomics
En. ludwigii LecVs2	*V. vinifera silvestris*	Endophytic	Genome sequencing, dual cultures, qPCr experiments
Erwinia sp. ErVv1	*V. vinifera vinifera*	Endophytic	Genome sequencing,
Escherichia coli DH5α	Invitrogen	Laboratory adapted	Colonization assays, Targeted metabolomics, sub-cloning
E. coli S17-1 λpir pMP4655	Bloemberg et al, 2001	Laboratory adapted	Propagation of plasmid pMP4655
Lysinibacillus fusiformis LyVs1	*V. vinifera silvestris*	Endophytic	Genome sequencing, dual cultures
Pantoea vagans PaVv1	*V. vinifera vinifera*	Endophytic	Genome sequencing, dual cultures, qPCr experiments
P. vagans PaVv7	*V. vinifera vinifera*	Endophytic	Genome sequencing, dual cultures, qPCr experiments
P. vagans PaVv9	*V. vinifera vinifera*	Endophytic	Genome sequencing
Pseudomonas chlororaphis 30-84	Pierson et al, 1994	Endophytic	Positive control quorum-sensing assays
Sphingomonas phyllosphaere SpVs6	*V. vinifera silvestris*	Endophytic	Genome sequencing, dual cultures
Botrytis cinerea	Fondazione Edmund Mach	Pathogenic	Biocontrol tests
Plasmopara viticola	(Musetti et al. 2006)	Pathogenic	Biocontrol tests

3.1.4.2. Plants

Figure 1. Maps of sampling sites across Italy where endophytes from wild (white circles) and domesticated grapevine (black elipse) were isolated. Color-climate legend: red - subtropical; orange - hot-temperate; light orange - sub litoranean; yellow - sub continental; light green - temperate; dark green - cold temperate; light blue - cold; dark blue – glacial.

In vitro micropropagated plants of *Vitis vinifera* L. cv Pinot noir (clone I-SMA 185) were prepared for experiments as described before (Compant et al. 2005). Briefly, the plants were micropropagated in cylindrical glass tubes on complete Murashige-Skoog (MS) medium pH 5.6 (Duchefa biochemie, The Netherlands) supplemented with 3% sucrose and 0.6% microagar (Duchefa biochemie, The Netherlands). This clone was chosen because in our collection it appeared free of bacterial contaminants and other bacterial endophytes, when tested in PCR using primers 799F/1520R as described elsewhere (Campisano et al. 2014a). Explants with one node and internode were incubated in a growth chamber for 51 days at 21°C, 16L: 8D h photoperiod and a photon irradiance of 50 µm s/m. Healthy plantlets with no less than 3 leaves and no signs of microbial contaminations were used for experiments. Greenhouse plants (*Vitis vinifera* L. cv. Pinot noir grafted on Kober 5BB) were grown under controlled conditions at 24 ± 1°C, 70 ± 10% relative humidity (RH) and a photoperiod of 16L:8D h. Plants were grown in pots on an organic plant substrate and were not treated with any pesticides for the entire course of the experiments.

3.1.4.3. Insects

Scaphoideus titanus eggs originated from two-year-old grapevine canes collected from organic farms in Northern Italy (Villazzano, Trento, Italy, 46°05'N, 11°14'E) during the first week of December 2014 and stored in a cool chamber (4±1°C). Starting from the beginning of April 2015, bundles of canes (0.5 kg) were weekly placed inside plastic boxes containing humid Perlite (Perlitech, Italy) in a climate chamber (24±1°C, 16L: 8D h photoperiod, 75% RH) where, after 30 to 60 days, eggs gradually hatched. Freshly hatched nymphs (IN) were removed daily and gently transferred to a SRC using a suction aspirator.

3.2. Biological functions and bioprospecting of bacterial endophytes from grapevine

3.2.1. Experimental design

All bacteria and fungi characterized in this study were isolated from grapevine endosphere as described above (see 3.1.3 and Figure 1). A total of 136 isolates from fiftysix *V. vinifera* subsp. *sylvestris* plants and 197 isolates from thirtytwo domesticated *V. vinifera* subsp. *vinifera* (cv Pinot Noir, Chardonnay and Merlot) plants were obtained. Wild grapevines were collected at different times during the year and grown for 4 years planted under the same environmental conditions in Northern Italy (Biagini et al., 2012). Domesticated plants were sampled in commercial vineyards in a comparable vine growing area in Northern Italy. Upon sampling,

all plant material was refrigerated and processed within 2 days from sampling. Only nine wild grapevine plants were originally collected in the North of Italy. As the collection areas only partially overlap, we calculated diversity indexes (Shannon index and Simpson index) for bacterial communities of wild grapevines in Italy and for the restricted group isolated in Northern Italy, to understand if the differences in sampling areas lead to overestimate the microbial diversity.

3.2.2. DNA barcoding by 16S rDNA sequencing

The 16S rDNA gene was amplified in all isolates, using the universal primers 16S-27F (AGAGTTTGATCCTGGCTCAG) and 16S-1492R (TACGGYTACCTTGTTACGACTT) and standard procedures (White et al. 1990). PCR-products, separated on a 1% agarose, were purified from the gel using Exo-SAP (Euroclone S.p.A., Italy) and sequenced using the BigDye terminator v3.1.

To identify the strains, DNA barcoding of the major endophytic populations was performed as follows: sequences were aligned against those from reference strains deposited in the National Centre for Biotechnology Information (NCBI) website and referenced in the List of Prokaryotic Names with Standing Nomenclature (LPSN) website and on literature (Euzéby 1997). Distance matrices were constructed using the p-distance as metric. The matrices were employed to reconstruct taxonomical dendrograms using the maximum-likelihood grouping algorithm with a bootstrapping of 1000 replications.

3.2.3. Bacterial automated ribosomal intergenic spacer analysis (B-ARISA)

To fingerprint bacterial endophytic communities, the Intergenic sequences (IGS) located between the small- and large-subunit rRNA genes were PCR-amplified. For bacterial ARISA (B-ARISA), DNA was extracted from surface-sterilized and aseptically peeled grapevine plants to avoid contamination from DNA of non-viable microorganisms from the surface. Vine sections adjacent to those used for isolation of endophytic microorganisms were used for DNA extraction. Stem discs were crushed using liquid nitrogen. DNA isolation, PCR, capillary electrophoresis, peak scoring and peak binning (the process of quantization resulting from grouping marker peaks into intervals) were performed as described elsewhere (Pancher et al. 2012).

3.2.4. Screening for enzymatic activity

Six tests for enzymatic activity screening were performed as follows. All tests based on agar plates (chitinase, cellulase, protease, lipase) were performed by inoculating appropriate plates (see Appendix Table 1 for details) with 5 µl of a bacterial cell suspension in SSS and left in incubation at 25°C for 5 days. Plates were daily checked for bacterial growth. Positive or negative reactions were scored (see Appendix Table 1 for conventions).

3.2.5. Screening for quorum sensing activity

The detection of quorum sensing signals was carried out using the bacterial biosensors *Agrobacterium tumefaciens* NT1 (pZLR4) and *Chromobacterium violaceum* CV026. *A. tumefaciens* NT1 (pZLR4) was used in petri dish assays according to Cha et al (1998) and *C.violaceum* was used plate T streak assay. Bacterial strains were scored as *N*-acyl homoserine lactone (AHL) producers if they restored violacein production in

strain CV026 or promoted lacZ transcription in strain NT1 (pZLR4). *P. chlororaphis* 30-84 and *E. coli* DH5α were used as positive and negative controls, respectively.

3.2.6. Screening of swimming & swarming

Five microlitres of a bacterial cell suspension in SSS were spotted in the centre of swimming or swarming plates (see Appendix Table 1). Media were incubated at 25°C for five days and checked for swimming and swarming (see Appendix Figure 1 for conventions).

3.2.7. Screening for biocontrol activity

Four tests were performed to establish the biocontrol potential of each endophyte.

3.2.7.1. Dual-plate antagonism

Co-cultures were done in petri dishes containing PDA (Oxoid, United Kingdom) against isolates of the grapevine pathogens *Botrytis cinerea*, *Botryosphaeria dothidea*, *Botryosphaeria obtusa*, *Phaeomoniella chlamydospora* and *Phaeoacremonium aleophilum*. Fungal pathogens were inoculated by placing 0.5 cm diameter agar plugs from the edge of an actively growing colony into agar plates. Bacterial endophytes were then inoculated in front of the fungal colony. Control plates were inoculated with the pathogen alone. The plates were incubated at 25°C, and the growth inhibition was scored when the control plate was fully covered by fungal growth. Strains were also tested for their biocontrol activity *in vivo* against *Botrytis cinerea* and *Plasmopara viticola*.

3.2.7.2. Biocontrol of Botrytis cinerea

Thousand grape berries (cv Victoria) were surface sterilized as described above. Each berry was stabbed in four different points using a sterile scalpel, and then sets of five berries were submerged in a suspension of bacteria (10^9 cfu/ml) and placed in a sterile box. After overnight incubation at room temperature, a suspension of 10^6 conidia/ml of *B. cinerea* was sprayed on the berries (20 ml each box). After seven days of incubation at room temperature, symptoms of grey mold were observed and scored for each of four wounds (each berry thus ranking 0 to 4) for each of the five barriers tested. Appropriate controls without biocontrol bacteria or without conidia from *B. cinerea* were performed. Constant high humidity in each box was provided by introducing sterile humidified tissue paper.

3.2.7.3. Biocontrol of Plasmapora viticola

This test was performed according to Pertot et al., 2006 with minimal changes. For each isolate, five leaf discs (2 cm diameter) were cut from greenhouse grapevine leaves, submerged in a bacterial cell suspension (10^9 cfu/ml) and then placed in a Petri dish. After overnight incubation at room temperature, a suspension of 4.25 x10^5 sporangia of *P. viticola* per millilitre of distilled water was sprayed on the leaf discs. The discs were incubated at room temperature for seven days. Inhibition by the tested strains was evaluated by comparing the presence or absence of sporulation of *P. viticola* on test leaf discs against the positive control (leaf discs inoculated only with sporangia of *P. viticola*). Negative controls were run with discs treated only with distilled

sterile water. Constant high humidity in each petri dish was provided as described above. Humidity was checked daily and tissue paper wetted when necessary.

3.2.8. Screening for plant growth promotion activity ex planta

Six tests were performed to establish the plant growth promotion of each endophyte, as described below.

3.2.8.1. 1-aminocyclopropane-1-carboxylate (ACC) - deaminase activity

Two hundred and fifty microliters of DF salt minimal medium (see Appendix Table 1 for details)) were plated on each of 48-well sterile plates and after solidification, 15 μl of a 50 mM ACC solution (Sigma-Aldrich, USA) were spread on each well. After complete drying of the ACC solution, 5 μl of a bacterial suspension were spotted on the wells. The plates were covered with a transparent plastic tape and incubated five days at 25°C.

3.2.8.2. Indole-3-acetic acid (IAA) production

Isolates were grown in DF salt minimal broth (see Table 1 for details) ammended with 500 μg/ml L-Tryptophan (Sigma-Aldrich, USA)for 2 days at 25°C. Two hundred and fifty microliters of grown bacteria in DF were mixed with 1 ml of Salkowski's ragent. After 30 min incubation at room temperature, results for IAA production were semi-quantitatively determined (see Appendix Table 1 for details).

3.2.8.3. Nitrogen fixation

Five microliters of a bacterial suspension in SSS were spotted on malate agar plates (16 strains each), and incubated 5 days at 25°C. A daily check of the plates was made. No source of nitrogen was present in the medium before bacterial growth (see Appendix Table 1 for details).

3.2.8.4. Phosphate solubilisation

NBRIP medium was inoculated with 5 μl of bacterial cell suspension in SSS (16 strains each), and incubated for 5 days at 25°C (see Appendix Table 1 for details).

3.2.8.5. Siderophore production

Chrome azurole S (CAS) agar was prepared (see Appendix Table 1 for details) and 5 μl of bacterial cell suspension in SSS were spotted on the plates (16 strains each), and incubated 5 days at 25°C.

3.2.8.6. Ammonia (NH_3) production

The testing strains were grown in ammonium medium (see Appendix Table 1) and incubated for 30min.

3.2.9. Screening for antibiotic resistance

Antibiotic resistance was screened on Petri dishes containing NA medium amended with the appropriate antibiotic as follows: Ampicillin (50 μg/mL); Gentamycin (10 μg/ml); Kanamycin (25μg/ml); Rifampicin (40 μg/ml); Streptomycin (15 μg/ml) and Tetracycline (10 μg/ml). The bacterial cell suspensions were spotted on plates (16 strains per plate) using a sterile rod, and the plates were incubated at 25°C in the dark, for 5 days. After incubation, complete inhibition of growth on the plates was scored as sensitive (score 0) to the correspondent antibiotic, colony size comparable to that formed after incubation on control NA plates for the same time was scored as resistant (score 2). When colony size was 50% or lower (but still visible to the naked eye) as compared to that on control NA plates, an intermediate score of 1 meaning partial resistance was assigned.

3.2.10. Growth promotion testing in planta

A subset of 40 strains was chosen among isolates with potential as growth promoters, according to the PGP aggregated score (with a score of 6 or higher, see Appendix Figure 1). The chosen bacterial strains were tested for their ability to promote plant growth as amendment to soil where *Arabidopsis thaliana* seeds were planted. To do this, cells of each bacterial strain were grown to early stationary phase, washed and dissolved in Hoagland solution (Hoagland and Arnon 1950). Sterilized soil was washed in Hoagland solution. Twenty grams of soil per replicate pot were mixed with about 10^6 cells of each bacterial strain. Twenty *A. thaliana* seeds (ecotype Col0) per replicate pot were surface-disinfected and planted on soil amended with bacteria. Five replicates were set for each test. Appropriate controls with non-autoclaved soil and autoclaved soil amended with sterile Hoagland solution were set. After 6 weeks, *A. thaliana* plantlets were harvested, counted and weighed. Bacteria inoculated in plots where the number of plantlets was in average 6 or higher (a number higher than that observed in the control where unsterilized soil was used), were considered strong promoters.

3.2.11. Statistical analysis

The PAST software (Hammer et al. 2001) was used for most of the analysis. A data Table was compiled, including all variables tested. Multivariate statistical analysis ws applied to the entire dataset produced, and to subsets of data, to infer information about the distribution of variance across data. Both taxonomic identity and origin of isolation were considered in the study. This, coupled with the use of multivariate analysis, permitted the dissection of differences among phenotypic traits in the endophytic microbial community. The data matrix was analyzed by principal component analysis (PCA).

Clustering of samples according to their origin of isolation (cultivated or wild grapevine, and in one case, wild grapevine from North, Centre, South or Insular Italy) was studied by one-way ANOSIM (Clarke 1993) and one-way NPMANOVA (Anderson 2001). These are non-parametric tests that analyse the significance of distances among multivariate groups.

3.3. Genome sequencing and analysis of bacterial endophytes

3.3.1. Experimental design

Genomes of seven bacterial endophytes, with biotechnological potential (namely *Enterobacter ludwigii*. EnVs6, *En*. *ludwigii*. EnVs2, *En. ludwigii* LecVs2, *Erwinia* sp. ErVv1, *Pantoea vagans* PaVv1, *P. vagans*. PaVv7, *P. vagans* PaVv9) belonging to the family Enterobacteriaceae, were sequenced. Twelve available genome sequences of reference strains, including human-, plant-pathogenic, endophytic or epiphytic strains, were selected for genome comparison and served as controls (*En. cloacae* subsp. *cloacae* ATCC 13047, *En. asburiae* LF7a, *En. aerogenes* KCTC 2190, *En.* sp. 638, *P. agglomerans* 299R, *P. ananatis* LMG 20103, *P. ananatis* PA13, *P. vagans* C9-1, *Er. billingiae* Eb661, *Er. amylovora* ATCC 49946, *Er. pyrifoliae* Ep1-96, *Er. pyrifoliae* Ejp617). The identity and genomic characteristics of these strains are summarized in Table 2 and in Appendix Table 2.

3.3.2. DNA extraction, genome sequencing and assembly

Test strains were grown in NB at 200 rpm on a SI600R rotatory incubator (MID SCI, USA) at 30±2°C until an OD_{600} = 0.8. Bacterial cells were pelleted on an ANNITA PK12 R bench centrifuge (ALC International, Italy) for 5 min at 10,000 rcf. The pellet was washed with sterile PBS 1X pH= 7,2 and DNA was extracted using the RBC real genomics DNA extraction kit (RBCBiosciences, China) according to manufacturer's instructions. DNA concentration was estimated using a Nanodrop 8000 UV-VIS spectrophotometer (Thermo Scientific, Germany). DNA integrity and the absence of RNA contamination were checked by electrophoresis on a 1% agarose gel. Sequencing libraries were constructed using the Nextera DNA Sample Prep Kit (Illumina, Inc., USA) according to the manufacturer's instructions. DNA-seq libraries were pooled at 10-plex level of multiplexing and sequenced paired-end 100bp on a HiSeq2000 Illumina sequencer at IGA Technology Services (Udine, Italy). Raw images were processed using Illumina Pipeline version 1.8.2.

The assembly of the genomes from test strains was performed as follows: three of the genomes (EnVs6, ErVv1 and PaVv9) were assembled using the A5 pipeline release 20140401 (Tritt et al., 2012); other four assemblies (PaVv7, EnVs2, LecVs2 and PaVv1) were produced using the SOAPdenovo software, version 2.04 (Luo et al., 2012). The assemblies were done using the two aforementioned pipelines as they provided the optimal number of contigs and N50, after testing different methods. Quality of the sequences was evaluated using the quality assessment tool for genome assemblies (QUAST) to produce metrics of quality using two different assembly pipelines (Alexey et al., 2013)

3.3.3. Annotation and subsystem analysis

Genomes were annotated using the Prokaryotic Genome Annotation System (PROKKA) (Seemann, 2014). To eliminate the bias of different annotation systems employed, test and reference genomes were all submitted to the online platform RAST (Rapid Annotation using Subsystem Technology) version 2.0. Six RAST subsystems central to our analysis (cell wall and capsule, iron acquisition and metabolism, chemotaxis, phages and mobilome, regulation and cell signalling, virulence and disease) were chosen to build a presence/absence map through a series of Perl® scripts (Wall, 2000), which summarize all pairwise comparisons into a unique list, including all genes assigned to the chosen subsystems. A list of genes

common to all compared genomes was compiled using a custom script in R version 3.0.2 (R Core Team, 2013)

3.3.4. Phylogeny

Sequences of the test strains' 16S rDNA genes were retrieved from RAST and confirmed that they were identical to those previously deposited in GenBank (Campisano et al., 2014). 16S rDNA sequences downloaded from RAST were then used to perform a blastn search against the NCBI database (Wheeler et al., 2007), and aligned with the closest sequences from the database using the clustalW algorithm implemented in the software bioedit version 7.2.5 (Hall, 1999). The phylogeny was reconstructed using the tree-building algorithm Neighbour-Joining with the Jukes-Cantor distance estimator implemented in MEGA6 version 6.0.6 (Tamura et al., 2013). Phylogeny, when possible, was assigned to the species level.

3.3.5. Whole-genome comparison

Several methods for whole-genome comparison were used to identify similarities and differences between each test and reference strain's genome in the same genus. First, the assembled genomes were retrieved from RAST and aligned against one reference genome of the same genus using MAUVE version 2.3.1 (Darling et al., 2004). Following this step a multiple whole–genome alignment was performed using the progressive alignment algorithm implemented in MAUVE. The output of this alignment was used to check for rearrangements in each genome.

Synteny plots (Husemann and Stoye, 2010) were constructed by aligning regions of the test and reference genomes that differed by no more than 8% and shared at least 44 overlapping 11mers no more distant from each other than 64 nucleotides. All regions were aligned and displayed in r2cat (Husemann and Stoye, 2010). Several plots were drawn against reference strains with different degrees of relatedness to check for synteny, for each of the three genera under study.

A circular genomic map was constructed for each genome using the BLAST Ring Image Generator (BRIG, version 0.95; Alikhan et al., 2011). Each circular genomic map was drawn using the genome of one reference strain (henceforth referred to as 'alignment reference genome') on a local BLAST+ basis, with standard parameters (50% lower – 70% upper cut-off for identity and E value of 10). For the genus *Enterobacter* the fully sequenced genome of the poplar endophyte *Enterobacter* sp. 638 was used as a reference. For the genus *Erwinia*, the sequenced genome of *Er. amylovora* ATCC 49946 was used. For *Pantoea* the genome sequence of the biocontrol agent *P. vagans* C9-1 was used. The ring colour gradients correspond to varying degrees of identity of BLAST matches. Circular genomic maps also include information on GC skew and GC content.

3.3.6. Clustering of orthologous families

In order to depict the core and accessory genome in each genus, a reciprocal best hit search using the OrthoMCL software release 5 was performed (Li et al., 2003). For this, the predicted coding sequences (CDS) of the reference and the test strains were downloaded and a blastp search against each other with an E value cut-off of 10^{-5} and a sequence coverage higher than 50%, was performed as previously reported (Li et al., 2013). A series of built-in scripts were used to i) parse the sequences, ii) upload them to the MySQL

relational database, iii) perform a reciprocal best hit analysis to form pairs of sequences, and to iv) normalize the E values for all the pairs formed. Normalization of E values was done by averaging all recent orthologues (in paralogues) and dividing each pair of orthologues by the average. Finally, using the Markov Cluster Algorithm (MCA) program, the sequences were distributed in orthologous families using 1.5 as the inflation value. The pangenome size was calculated as the sum of the CDS in all the genomes and the accessory genome size (within each genus and for each species) was calculated as the difference between the pangenome and the core genome. Functions to all orthologous families were assigned and virulence functions were filtered using a set of Perl® scripts. A list of keywords was also used to query for virulence functions in the orthologous families and to calculate the number of matches of those functions, using custom bash commands**.**

3.3.7. Secretion system analysis

Two databases (the Lawrence Livermore national laboratory virulence database -MvirDB; Zhou et al., 2007- and the virulence factors of pathogenic bacteria database -VFDB; Chen et al., 2012) were used to download datasets of virulence genes (either validated or predicted). To be mapped on the sequenced genomes. The matching results were drawn in a presence/absence map.

3.3.8. CRISPRs and phage presence

To evaluate the presence of CRISPRs (Clustered Regularly Interspaced Short Palindromic Repeats) the assembled genomes of the seven test strains were analyzed using the CRISPR Recognition Tool CRT (Bland et al., 2007). Also a search for putative phage sequences in the test strains in the online tool PHAST (Zhou et al., 2011) was performed.

3.4. Transmission of bacterial endophytes using Scaphoideus titanus as vector

3.4.1. Experimental design

Four two-year-old greenhouse grapevine plants were grown as described above (see 3.1.3.1). These plants are hereafter referred to as “source” (SRC), since they host the typical complex microbial community of plants grown under natural conditions (Campisano et al., 2014a).

A total of 35 *in vitro* micropropagated grapevine plantlets were prepared as described above (see 3.1.3.2). These *in vitro* plantlets are hereafter referred to as “sink” plants (SNK) and they represent the plants where the bacterial community will be delivered. To further exclude any bacterial presence in the tissues, ten of these 35 SNK were used as controls.

Scaphoideus titanus individuals were prepared as described above (see 3.1.3.3). Freshly hatched nymphs (IN) were removed daily and gently transferred to a SRC using a suction aspirator. The transmission experiment (Figure 2 left) was carried out independently four times, using new plants and insects.

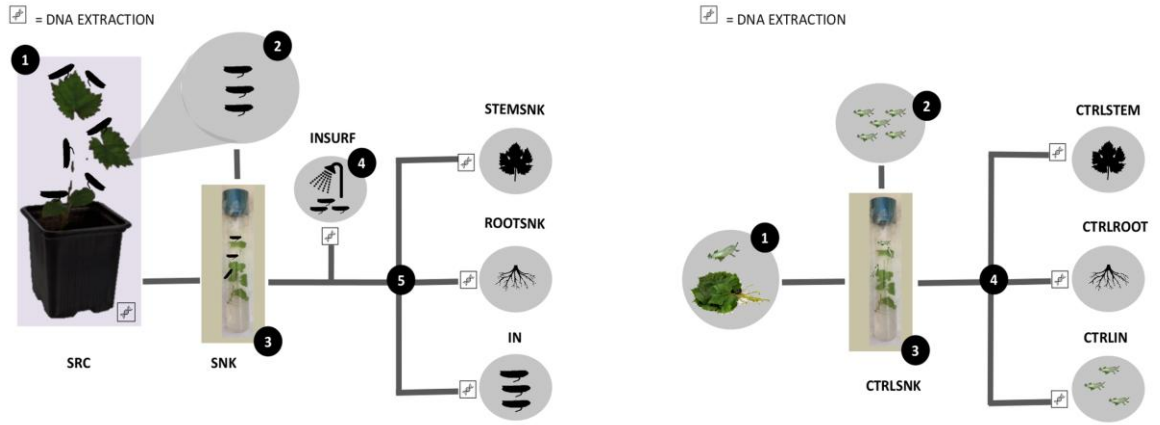


Figure 2. Experimental design of artificial transmission using *S. titanus* as vector. Left panel- SRC: Source plants; SNK: Sink plants; INSURF: Insect surface; STEMSNK: stems of sink plants; ROOTSNK: roots of sink plants; IN: Insects; Right panel - CTRLSINK: control sink plants (not in contact with the source plant-microbiota); CTRLSTEM: Stems of control plants; CTRLROOT: roots of control plants; CTRLIN: Control insects (not in contact with the source plants)

Four SRC were kept under constant environmental conditions as mentioned above. Then, 96 IN were placed and confined onto four well developed SRC leaves, where restricted areas were delimited by small cages (four cages per plant, with six INs each for a total of 24 insects per plant) made out of a mesh sleeve (250 µm mesh size) and a supporting plastic cylindrical structure (ø = 10 cm; h = 20 cm). The IN were let to feed and grow for 14 days until they reached a stage between the third and fourth nymphal instar. Then, out of the 96 IN previously transferred, 48 individuals were collected from the SRC (12 IN per SRC) and transferred to 16 SNK (four SNK with three IN per each SRC). In addition, an insect-free SNK per each SRC was included as a sterility control of the replicate. Before transferring the IN, surface of the MS medium supporting the micropropagated SNK was overlaid with 1 ml of sterile melted paraffin (Sigma-Aldrich, Germany) in order to prevent the contamination of the growth medium and the roots by microorganisms carried by IN. In this way, we could assure that no contact between the roots or the growth medium was taking place. Once transferred to the SNK, IN were allowed to feed for 10 days at 21°C, 16L:8D h photoperiod and a photon irradiance of 50 µm s/m^2 until they were fifth instar nymphs or adults.

As control, five SNK (CTRLSNK) were each infested with five freshly hatched nymphs (CTRLIN) that had not been previously reared on SRC, but were feeding only on SNK (Fig. 2, right). In addition, from the remaining ten SNK, five were used to probe for bacterial DNA in the plant's tissues. Total plant DNA was extracted using the method previously described (Campisano et al., 2014a) and the extracted DNA was amplified using the primer pair 799F/1520R (Yousaf et al., 2014). Since the five tested plants were PCR negative (no amplification of bacterial 16SrDNA gene), they were considered bacteria-free. Although primers 799F and 1520R are a universal pair for 16SrDNA amplification, still some prokaryotes might not have been detected. Thus, the last five SNK plantlets were used to control microbial contamination inside the tissues. SNK were incubated under the same conditions without IN. Then, plants were crushed in a sterile mortar with 1 ml

phosphate buffer saline 1X, pH 7.2, and the resulting extract was plated on Luria-Bertani agar (LBA; Sigma Aldrich, Germany) and incubated at 30°C for five days, after which no growth was recorded.
After the incubation period, all SNK and IN were aseptically removed from the glass tubes. SNK were cut into stems (STEMSNK) and roots (ROOTSNK); CTRLSNK were likewise cut into CTRLROOT and CTRLSTEM samples. IN were washed with distilled sterile water by thoroughly vortexing in order to dislodge the majority of surface-adhering bacteria. The bacterial cells in the washing water (INSURF) were pelleted by centrifugation at 13,000 rpm on a Tabletop centrifuge and stored at -20°C before extracting the DNA. All SRC, IN, CTRLIN, STEMSNK, ROOTSNK, CTRLSTEM and CTRLROOT were then surface-sterilized (see 3.1.2)

3.4.2. DNA extraction, 16SrDNA amplification and pyrosequencing

After sterilization, SRC, IN, CTRLIN, STEMSNK, ROOTSNK, CTRLSTEM and CTRLROOT were aseptically transferred to sterile stainless steel capsules containing steel beads. The material was frozen in liquid nitrogen for 5 min and crushed in a Retsch MM200 tissue lyser (Qiagen, The Netherlands) for 2 min at a frequency of 25 hertz. The resulting powder was weighted and then deoxyribonucleic acids were extracted using the FastDNA™ SPIN Kit for Soil (MP, United States) according to manufacturer's instructions. DNA from the INSURF samples was extracted with the same kit after pelleting and suspending the cells in extraction buffer before workup.
DNA was then quantified in an UV-VIS nanodrop 8000 spectrophotometer (Thermo Fischer Scientific, United States) and PCR-amplified using the primer pair 799F (AACMGGATTAGATACCCK) and 1520R (AAGGAGGTGATCCAGCCGCA) targeting the V5 – V9 16S rDNA hypervariable regions without amplification of plastid DNA. These primers bear 454 adaptors and a sample-specific barcode on the forward primer. PCR was performed using the Roche high fidelity Fast Start PCR system (Roche, Switzerland) in a final volume of 25 µl. The following volumes, reagents and concentrations were used: 2.5 µl amplification buffer 10X, 5 µl $MgCl_2$ 25 mM, 0.5 µl reverse primer 10 µM, 0.5 µl forward primer 10 µM, 2.5 µl dNTPs 25 mM, 1 µl DMSO, 2.5 Bovine serum albumin (BSA) 10 mg/ml, 0.4 µl HI-FI Taq polymerase 5U/µl and water. DNA was adjusted to an initial concentration of 3 ng/µl and for some samples dilutions of 1:10 were used in order to obtain optimal amplification. Thirty cycles of PCR were carried out according to the manufacturer's instructions with conditions for amplification as follows: 5 min of initial denaturation at 95°C, 30 s at 95°C, 1 min for annealing at 53°C, 2 min for extension at 72°C, and a final extension step 10 min at 72°C. PCR products were separated in a 1.5% agarose gel stained with SYBR® Safe DNA Gel Stain (Thermo Fisher Scientific, United States), and visualized on a Gel Doc XR+ system (BiO-RAD, United States). The appropriate amplification bands were excised from the gel using the PureLink Quick gel extraction Kit (Thermo Fisher Scientific, United States) according to manufacturer's instructions. Three different amplifications for each sample were performed, purified from gel and pooled together for pyrosequencing. Amplicons were quantified with quantitative PCR using the library quantification kit Roche 454 Titanium (KAPA Biosystems, United States) and pooled in equimolar ratio in the final amplicon library. Pyrosequencing was carried out on the Roche GS FLX+ system using the new XL+ chemistry dedicated to long reads of up to 800 bp, following the manufacturer's recommendations.

3.4.3. Bacterial 16SrDNA amplicon demultiplexing and statistical analysis

Outputs from the 454 pyrosequencing were analyzed using the "Quantitative Insights into Microbial Ecology (QIIME)" pipeline, version 1.9.0 (Caporaso et al., 2010b). The analysis consisted of decoding the sequence flowgram files (SFF) and producing fasta and quality files with which length of sequences and quality of reads were checked. Amplicon sequences were demultiplexed (assigned to sample pools) according to their barcoded primer. Only bacterial sequences at least 300 nucleotides (nt) long were retained. Sequences were truncated when the quality score in a 50 nt long sliding window went below 25.

Chimeric PCR products were identified using USEARCH 6.1.544 (Edgar, 2010). Operational taxonomic units (OTUs) were picked using a threshold identity of 97% and the Greengenes database, August 2013 version (DeSantis et al., 2006). USEARCH cluster seeds were used as representatives for OTUs, while taxonomy was assigned using USEARCH and the Greengenes database as a template. Sequences assigned to chloroplasts and mitochondria were removed. OTUs represented by only one or two reads (singletons and doubletons) were removed from the OTU Tables. The amplicons were then aligned *de novo* using pynast (Caporaso et al., 2010a) and the alignment was used to generate a phylogenetic tree.

From pyrosequencing we obtained 1,404,963 reads from the whole set of samples, with a median of 13,271 reads per sample. After the first quality control steps where we removed short sequences (less than 200 nt), mis-sequenced fragments and mutated amplicons, only 1,024,657 sequences were left. Following removal of chimeric sequences using the Usearch algorithm, 871,497 remained as non-chimeric sequences. Here, a maximum of 31,039 sequences for IN samples and a minimum of 57 sequences for CTRLIN samples were obtained, and a mean of 12,939 sequences for all the samples.

For clustering OTUs, we picked a representative set of sequences that further represented the OTU with 97% accuracy, resulting in 2,005 grouped sequences available for analysis. Some of the sequences obtained were found to be of plant nature (plastid sequences) and were removed, leading to a final count of 1,923 sequences. From those, we removed the sequences that were represented in less than 1% of the total population, obtaining an OTU Table with a total of 447 OTUs that were defined as clusters composed of three or more sequences.

Alpha- and beta-diversity were estimated on multiple OTU Tables rarefied to 1,300 reads (considering the sample with the lowest number of reads). Alpha-diversity differences were tested for statistical significance using 999 Monte Carlo permutations and the p-value obtained corrected using the Bonferroni correction for multiple comparisons. Beta-diversity was computed using the phylogenetic unweighted UniFrac distances. PCoA plots rendering sample distances were visualized using Emperor and further drawn in R. A Kruskal-Wallis test was used to assess if the differential distribution of OTUs and taxa was statistically significant for all the variables analyzed. The multivariate test ANOSIM to detect differences between groups of samples was used as implemented in QIIME.

3.4.4. Transmission and quantification of endophytes through qPCR

To quantify bacteria transferred by *S. titanus* across plants, a similar setting to the one described in 3.4.1 was used. In this case, the source of inoculum was not the SRC, but a bacterial cell suspension of cultivable endophytes isolated from grapevine trunks (see 3.2). These bacteria were classified as *Enterobacter ludwigii* EnVs6, *E. ludwigii* EnVs2, and *Pantoea vagans* PaVv9 (see 3.2)

These bacterial endophytes were transformed with the eGFP-encoding plasmid pMP4655 (Bloemberg et al., 2000) as follows: cells were grown on LBA for 48 h at 30ºC. Then, 2 ml of super optimal broth amended with sucrose (SOC) were inoculated with a single colony and incubated for 24 h at 30ºC and 160 rpm (Hanahan, 1983). Aliquots of 400 µl of this starter culture were inoculated into 40 ml of SOC broth and then incubated for further 24 h at 30ºC and 160 rpm. Cells were then centrifuged at 45,895 rpm for 15 min at 4ºC, and subsequently suspended in electroporation buffer (EB) for plasmid insertion into the bacterial cells. Three washing steps were performed with EB, reducing in halves the resuspension volume. At the end, aliquots of 50 µl of buffered immersed (competent) bacteria were dispensed in tubes and kept at -80ºC. Bacterial cells were then gently mixed with 1 µg of plasmid and incubated on ice for 30 min. Later, the mixture was transferred to 0.2 cm electroporation cuvettes (Biorad, United States) and electroporated at 1500 mV, 25 µF and 200 Ω. Cells were immediately immersed in 800 µl of SOC and incubated at 30ºC and 160 rpm for 2 h. Cultures were centrifuged and half of the volume discarded. Then, cells were suspended in the remaining volume and plated onto LBA supplemented with tetracycline (20 µg/ml). Transformants were confirmed by amplifying the resistance marker cassette tetA/R, present in the plasmid, with primers directed towards the gene, as previously reported (Møller et al., 2016).

Endophytic cells bearing the pMP4655 were grown on LB for 24 h and cell densities were adjusted to 3×10^7 CFU/ml. Then, cells were cooled down on ice and washed three times with PBS 1X, pH 7.2. After the last washing step, cells were re-suspended in 200 µl of a Tris-EDTA-sucrose pH 8.0 solution (TES: Tris 10 mM, EDTA 1mM, sucrose 5% w/v) and distributed in the lids of bottomless (replaced by a cotton plug) 1.5 ml plastic tubes (Eppendorf , Germany). Lids were covered with one layer of sterile parafilm (Bemis NA, United States).

Scaphoideus titanus individuals were reared as described above (see 3.1.3.3). Insects were transferred to plastic tubes with the lids hanging upside down, and left to feed on the eGFP-tagged bacteria for 5 days. In this setting, insects punched the parafilm layer on the lid, releasing and feeding from the bacteria-rich TES solution. After feeding, insects were transferred to *in vitro* micropropagated grapevine plantlets, as described in the experiments above. Each replicate consisted of four plants infested with three insects per plant and incubated for five days. At the end of the incubation, the insects, roots and stems were collected separately. DNA from the insects and plants was extracted using the NucleoSpin® Plant II kit (for ROOTNSK, STEMSNK, CTRLROOT and CTRLSTEM) and the Nucleospin® Tissue (for IN and CTRLIN) according to manufacturer instructions (Macherey-Nagel, Germany). DNA was quantified as described above.

Bacterial DNA, including the pMP4655 eGFP encoding plasmid, was quantified on a Roche LightCycler® 480 Real-Time PCR (Roche, Switzerland) with the platinum SYBR Green qPCR superMix-UDG (Thermo Fisher Scientific, United States). The following amplification protocol was used: 1 hold at 50ºC activation (UDG incubation) for 5 min, 1 hold at 95ºC for 5 min activation, 40 cycles at 95ºC for 30 s for melting and at 58ºC for 45 s for annealing and extension. An analysis of melting curves at 95ºC for 5 s, followed by a cooling down until 55ºC for 1 min, was performed to check for specificity of the reaction. Absolute quantification of eGFP gene copies in plants and insects was done based on interpolation from a standard curve obtained with serial 10-fold dilutions of the eGFP gene (from 3×10^6 to 3×10^1 eGFP gene copies/µl) in DNA of control plants or insects, respectively.

3.5. Colonization assays using bacterial endophytes from grapevine

3.5.1. Experimental design

In vitro micropropagated *Vitis vinifera* plantlets were prepared for inoculation as described above (see 3.1.4.2). Plants were transferred to sterile plastic boxes containing 40 ml of MS agar inoculated with 100 µl of a bacterial cell suspension at a concentration of $3x10^8$ CFU/ml (corresponding to the OD_{600} value of 0.1). Plants were then kept in the inoculation chamber and incubated for 10 days using the same photoperiod and temperature conditions described above.

Further twenty four micropropagated plantlets of *Vitis vinifera* were cultured as described above. Three replicates of four plantlets each were inoculated with strain *E. ludwigii* EnVs6 and three replicates consisting of four plants each were inoculated with *E. coli* strain DH5α (a non-endophytic, non-pathogenic laboratory strain that served as control) and kept in growth chambers under the same incubation conditions adopted for plants used for microscopic observation of tissue colonization.

After 10 days from inoculation, all plantlets were frozen in liquid nitrogen, crushed in a Retsch MM200 tissue lyser (Qiagen, The Netherlands) for 2 min in screw-cap steel capsules containing steel beads, at a frequency of 25 herz. Finally, the four grapevine plantlets that formed a biological replicate were pooled. In a second replicate experiment, the identical procedure as described above was followed, but below- and above-ground plant organs were aseptically separated before the freezing step, with the purpose of confirming the distribution of secondary metabolites between plant organs.

3.5.2. Double labeling of oligonucleotide probes-Fluorescence in situ hybridization (DOPE-FISH)

DOPE-FISH was performed on bacterial pure cultures alone and on plants inoculated with bacteria. For axenic cultures, bacteria were inoculated in LB medium and incubated at 27°C and 120 rpm on an orbital shaker until the exponential growth phase. Cells were harvested by centrifugation at 4500 x g for ten minutes and washed several times with PBS 1X pH 7.2. Following washing, cells were fixed in a 4% v/v paraformaldehyde (in PBS) solution at 4°C overnight and then treated with a lysozyme solution (1mg/ml) for 10 minutes at 37°C. Cells were then rinsed three times with PBS and centrifuged at 4500 x g for 10 minutes in every washing step. Later cells were dehydrated in increasing concentrations of ethanol solutions (25, 50, 75 and 99%), and stored at 4°C until further use. Cells were poured into teflon-coated microscope slides (Immuno-cell, Belgium), air dried and hybridized according to Compant et al (2005) using probes EUB338, EUB338II, EUB338III (EUBmix) labelled with FLUOS and Gam42a labeled with Cy5 (Amann et al. 1990; Daims et al. 1999; Manz et al. 1992; Wallner et al. 1993) with fluorochromes at both 5' and 3' ends. NONEUB coupled with Cy5 was used as a control of the experiments. Hybridization step was carried out with 20 µl of hybridization buffer (NaCl 0.9 M; Tris-HCl 0.02 M; 0.01% SDS, 35% formamide and probes at a concentration of 5 ng/µL) and slides were placed in 50 ml moisture chambers filled with 5 ml hybridization buffer. Hybridization was carried out for 2 hours at 46°C in the dark followed by a post-hybridization step at 48°C during 30 minutes using a pre-warmed solution (20mM Tris-HCl pH 8.0; 0.01% SDS; 5mM EDTA and NaCl corresponding to the formamide concentration used). Samples were then rinsed with distilled water and air dried overnight in the dark.

To observe colonization by endophytic bacteria, plantlets were aseptically dissected into roots, stems and leaves as described by Compant et al. (2005). Samples were then cut in small parts (5 mm), fixed and prepared for DOPE-FISH as described above for bacterial cells. Then the plant material was transferred to teflon-coated microscope slides (Immuno-cell, Belgium) and hybridized with probes EUB Mix and Gam42a as described above. Some samples were sectioned transversally using razor blades. The hybridization and post hybridization were carried out as described above for bacterial cells and five replicate plants were analyzed for each strain under study. Another five replicate plants inoculated with sterile PBS 1X pH 7.2 were prepared as control as well. Finally, five plants per strain were used to test probe specificity with the NONEUB probe. Samples were rinsed with distilled water before being air dried overnight in the dark and analyzed under confocal microscope (Olympus Fluoview FV1000 with multi-line laser FV5-LAMAR-2 HeNe(G)laser FV10-LAHEG230-2). Pictures were taken at 405, 488, 633 nm wavelengths and under normal light and then merged (RGB) using image J software. Pictures were also analyzed using Imaris 8 software (BITPLANE, UK). Z-stacks were then used to generate whole-stack pictures, these pictures were sharpened (removing convolution by built-in microscope software), and the light/contrast balance was adjusted to improve detail visualization as seen when samples were observed in the dark conditions under the microscope.

3.5.3. Metabolic profiling of colonized grapevine plants

Crushed plant material (see 3.5.1) was analyzed according to previously established methods (Vrhovsek et al. 2012). Briefly, 0.1 g of crushed material were extracted in 2 ml Eppendorf tubes with 5 ml of a water/methanol/chloroform (1:2:2) mixture. Additionally, 20 µl of internal standards (gentisic and rosmarinic acids 50mg/l) were added. Samples were mixed by vortexing for 1 min and incubated in an orbital shaker for 15 min at room temperature. Samples were centrifuged at 15000 x g at 4°C for 5 min, and the aqueous phase was collected. Extraction from the pellet was repeated using 600 µl of water/methanol (1:2) and 400 µl of chloroform, by shaking for 15 min. After centrifugation, the two aqueous phases were pooled, dried under a nitrogen stream and dissolved in 500 µl of methanol/water (2:1). Samples were transferred to glass vials and stored at -20°C before injection.

Ultraperformance liquid chromatography was performed on a Waters Acquity UPLC system (Milford, USA) consisting of a binary pump, an online vacuum degasser, an autosampler, and a column compartment. Separation of the phenolic compounds was achieved on a Waters Acquity HSS T3 column 1.8 µm, 100 mm × 2.1 mm (Milford, USA), kept at 40 °C. The mobile phase A was water containing 0.1% formic acid, the mobile phase B was acetonitrile containing 0.1% formic acid. The flow was 0.4 ml/minute, and the gradient profile was: 0 minutes, 5% B; from 0 to 3 minutes, linear gradient to 20% B; from 3 to 4.3 minutes, isocratic 20% B; from 4.3 to 9 minutes, linear gradient to 45% B; from 9 to 11 minutes, linear gradient to 100% B; from 11 to 13 minutes, wash at 100% B; from 13.01 to 15 minutes, back to the initial conditions of 5% B. The injection volume of both the standard solutions and the samples was 2 µl. After each injection, the needle was rinsed with 600 µl of weak wash solution (water/methanol 90:10) and 200 µl of strong wash solution (methanol/water 90:10). Mass spectrometry detection was performed on a Waters Xevo TQMS (Milford, USA) instrument equipped with an electrospray (ESI) source. Analysis was done in positive and negative polarities. Flow injections of each individual metabolite were used to optimize the MRM conditions.

3.5.4. Statistical analysis

Quantitative results of UPLC-MS analysis of metabolites were studied by univariate and multivariate methods. PCA was performed by PAST 3.05 software (Hammer et al. 2001) on data from experiments one and two independently and on the entire dataset. Raw data were transformed to row percentages before analysis.

To find differences in the concentration of particular metabolites in plantlets treated with strain EnVs6, multiple t-tests were performed by correcting for false discovery rate (FDR) at q=0.01, using PRISM graphpad software version 6. The two experimental replicates were analyzed together.

A two-way ANOVA was performed to determine the effect of treatment in both experiments, using metabolite and treatment as factors, at α=0.05.

To reveal the effects of plant organ and treatment on the concentration of each metabolite, a generalized linear model (GLM) was implemented, in which the response variable was the concentration of each metabolite and the model tested differences for the treatment, organ and the interaction organ*treatment at

4. RESULTS AND REMARKS

4.1. Diversity in endophyte populations reveals functional and taxonomic diversity between wild and domesticated grapevines

The first step in our investigation was to establish the potential use of endophytes from grapevine, and establish if they actually hold beneficial properties for the plant, as has been shown for many other models of endophytism. Understanding these properties and correlating them to their ecology (do the properties change when endophytes inhabit a different host? Has domestication of grapevine affected the biotechnological potential of these microorganisms? Which are the most represented taxa in grapevine´s endophytic microbiota?) will shed light into the meaning of ecological interactions of endophytes with their host, in the context of bioprospecting.

A collection of 333 bacterial endophytes was obtained from surface sterilised grapevine stems (197 from domesticated grapevine and 136 from wild grapevine). One hundred and fifty-five unique strains (37 from domesticated grapevine and 118 from wild grapevine) were grouped into 28 genera on the basis of their taxonomic identity and used for this study. Bacterial isolates from domesticated grapevines were grouped in six genera only, while isolates from wild grapevines were grouped in 27 genera. The most abundant genera were *Pseudomonas* (24 strains), *Pantoea* (22 strains), *Bacillus* (20 strains), *Microbacterium* (15 strains), *Sphingomonas* (14 strains) and *Curtobacterium* (10 strains). Only one bacterium, which was isolated from domesticated grapevines alone, belonged to the genus *Erwinia*. Shannon's diversity index of bacterial endophytic populations was 1.424 for domesticated and 2.815 for wild grapevine. The Simpson index was was 0.5465 for domesticated and 0.1429 for wild grapevine.

4.1.1. Endophytic bacterial communities in wild and cultivated grapevine differ in diversity

Bacterial ARISA resulted in 267 markers of bacterial communities in wild grapevine (430 before peak binning) and 44 markers in domesticated grapevine (101 before peak binning), indicating that a broader diversity of endophytic bacteria inhabits wild grapevines than domesticated ones (Appendix Figure 2).

All unique strains were screened for plant growth promotion (PGP), resistance to antibiotics (AR), production of lytic enzymes and competition for nutrients (ENZ), quorum-sensing related traits such as cell motility and *N*-AHL production (QS) and biocontrol activity (BiCo). All results are reported in Appendix Figure 1 and summarised in Figure 3.

By analysing the generated scatter plots, we highlighted that: a) a PCA scatterplot of all endophytes according to the entire set of traits screened here (Figure 3a) roughly grouped them in three clusters; b) bacterial endophytes identified as genus *Pseudomonas* grouped in two clusters (well overlapping with the groups isolated from wild and domesticated grapevines, Figure 3b); c) similarly to *Pseudomonas*, the major bacterial endophyte groups (*Bacillus*, *Curtobacterium, Microbacterium* and *Pantoea*) clustered according to the host (wild or domesticated grapevines, Figures 3c,-f).

One-way ANOSIM (Clarke 1993) showed that that endophytes of domesticated grapevine were significantly different from those of wild grapevines (Bonferroni-corrected p value < 0.001, using Euclidean distances. This difference could be also observed when comparing score relative to PGP, QS and BiCo.

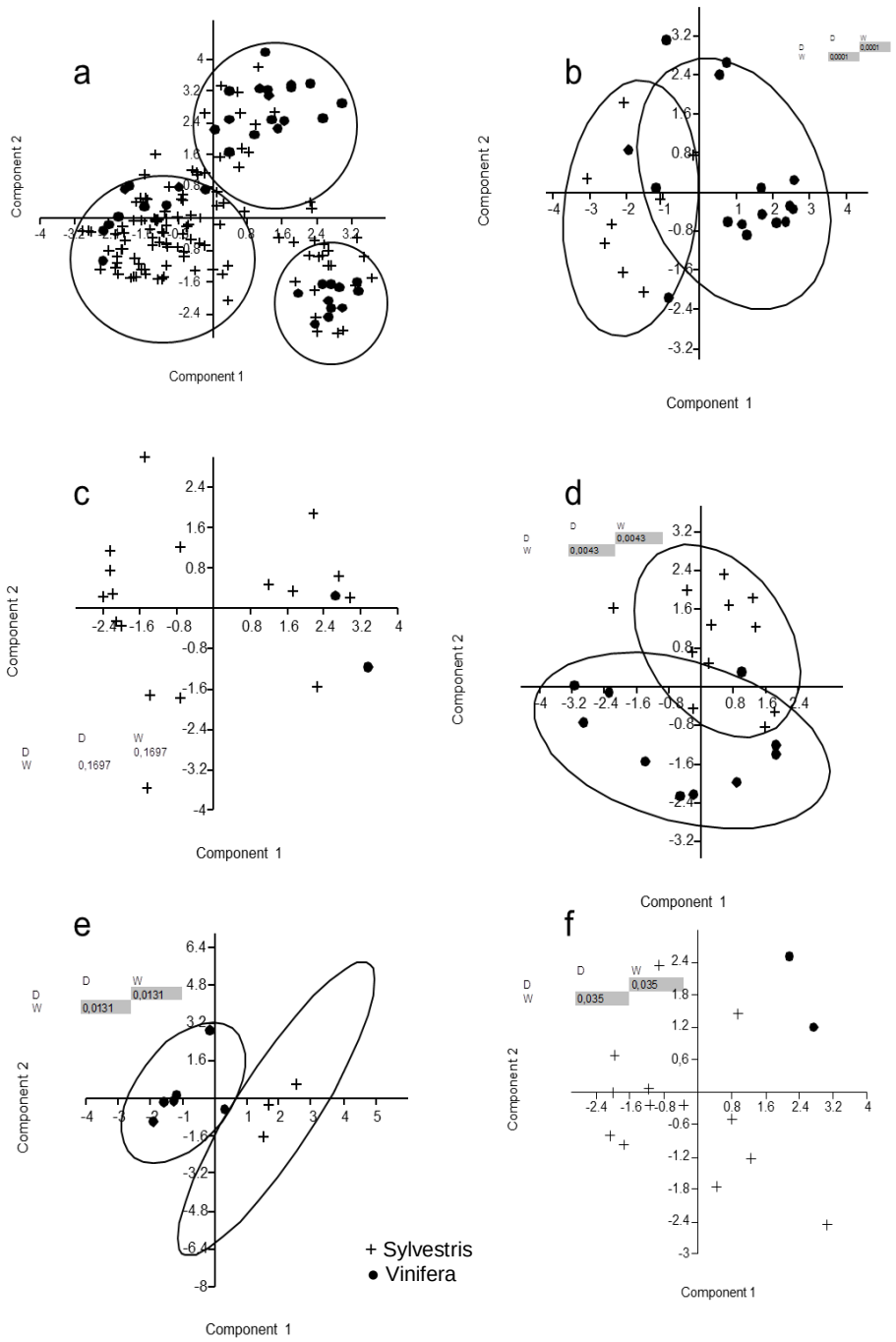


Figure 3. Principal component analysis of biotechnological properties of endophytes from grapevine. a) PCA scatterplot of all endophytes tested in this study; b) PCA scatterplot of *Pseudomonas* strains; c) PCA scatterplot of *Bacillus* strains; e) PCA scatterplot of *Curtobacterium* strains; f) PCA scatterplot of *Microbacterium* strains. Ellipses are designed using a confidence threshold of 75%. The inset tables ables represent p values of one-way ANOSIM test.

4.1.2. Endopytes are a rich source of enzymatics functions

4.1.2.1. Growth promotion ex planta

Clear ACC deaminase activity was detected in 56 strains. Nineteen strains among the ACC deaminase positive, were isolated from domesticated grapevine. ACC deaminase activity was detected isolates afiliated to *Enterobacter* spp. *Brevundimonas* spp., and *Pantoea* spp. In 9 strains isolated from wild grapevines, IAA production was high (score 4 on a 0-4 scale). These strains were assigned to *Sphingomonas*, *Pseudomonas*, *Paracoccus*, *Microbacterium* and *Bacillus*. Medium-high levels (score 3) of IAA were detected in 6 strains (5 isolated from domesticated grapevine and 1 from wild grapevine), 19 strains produced medium levels (score 2) of IAA. Out of 34 strains producing medium to high level of IAA (scoring 2 or higher on our scale) 15 were isolated from domesticated grapevine and 13 from wild grapevine. Nitrogen fixing ability was evident in the majority of tested strains. The ability to solubilise phosphate was limited to 36 strains. Of these, 16 were isolated from domesticated grapevine. Among the endophytes with a PGP score of 9 or higher, those isolated from domesticated grapevines were significantly overrepresented (12 out of 17, $p= 6.1x10^{-5}$ using an exact binomial test of goodness-of-fit).

4.1.2.2. Tests for antibiotic resistance

Only 48 strains were sensitive to ampicillin, 42 were partially resistant and grew slowly, 67 were fully resistant to ampicillin. Gentamycin-supplemented NA medium fully supported the growth of 19 strains; the growth of 30 strains was slowed down, while the growth of 93 strains was completely inhibited. Twenty-eight strains were fully resistant to kanamycin, 36 were partially resistant, and 78 were sensitive. Thirty-seven strains were fully resistant to rifampicin, 31 were partially resistant, and 74 were sensitive. Thirty-four strains were fully resistant to streptomycin, 25 were partially resistant, and 83 were sensitive. Only 17 strains were fully resistant to tetracycline, 29 were partially resistant, 96 were sensitive. An overall antibiotic resistance index (namely AR) was obtained by adding up the scores of each of the 6 tests (thus ranging 0 to 12). The best scoring on the aggregate AR index was strain 21b identified as *Sphingomonas* sp. (AR score 12). Six strains scored 10 or higher in the aggregate AR index, interestingly 4 out of 6 belonged to the genus Pantoea.

4.1.2.3. Enzyme production and competition for nutrients

Chitinase activity was detected only in one strain, identified as *Brevundimonas* sp. Fourteen genera and 45 strains tested positive for protease activity, among these 9 were *Pantoea*, 7 *Bacillus*, 7 *Pseudomonas* and 5 *Enterobacter*. Lipase activity was detected in 5 strains, belonging to 5 different genera: *Massilia*, *Microbacterium*, *Pantoea*, *Paracoccus* and *Rhizobium*. Cellulase activity was detected in 41 strains (belonging to 15 genera). The taxa that exhibited most of the cellulose activity were *Bacillus*, *Sphingomonas* (6 strains each), *Enterobacter*, *Microbacterium* and *Pantoea* (5 strains each). Phosphatase activity was not detected in only 31 endophytic isolates. Twenty-five strains exhibited low levels of phosphatase activity. Galactosidase acticity was detected in 37 strains. The most represented taxa among those with galactosidase activity were *Pantoea* (10 strains) and *Pseudomonas* (9 strains). An index summarising all

tests in this area was calculated by adding up all individual test scores (namely ENZ). The best ENZ score was observed in a strain identified as *Enterobacter ludwigii*.

4.1.2.4. Biocontrol activity.

Twenty-five endophytic strains were very active against *B. cinerea in vitro* (with score 2), among these, the most frequent were those belonging to genera *Bacillus* and *Pantoea* (5 strains of each genus). Thirteen strains exhibited moderate activity (score 1) against *B. cinerea in vitro*. Twenty-four strains were very active against *B. dothidea* in plate essays, among these, the most frequent were those belonging to *Bacillus* and *Pantoea* (5 strains of each genus). Nine strains exhibited moderate activity against *B. dothidea*. Twenty-eight strains were very active against *B. obtusa* in plate assays. Among these, the most frequent genus was *Pantoea* (6 strains). Nine strains exhibited moderate activity against *B. obtusa*. Forty-four strains were active against *P. chlamydospora*, the most frequent were those belonging to genera *Pseudomonas* and *Pantoea* (14 and 13 strains respectively). Strains isolated from domesticated grapevine were significantly overrepresented (26 out of 44, $p= 9.7x10^{-7}$ using an exact binomial test of goodness-of-fit) among those active against *P. chlamydospora*. Only 6 strains were very active against *P. aleophilum*, and 7 strains showed moderate biocontrol activity. When tested *in vivo* against *B. cinerea*, 27 strains were very active against this pathogen (score higher that 1.5 on a 0-2 scale), the most frequent were those belonging to genus *Pantoea* (10 strains). Strains isolated from domesticated grapevine were significantly overrepresented (14 out of 27, $p= 2.4x10^{-3}$ using an exact binomial test of goodness-of-fit) among those active *in vivo* against *B. cinerea*. Forty-one strains were very active against *P. viticola* infection. Strains active against *P. viticola* belonged to 16 genera. The most represented genus was *Pseudomonas* (8 strains).

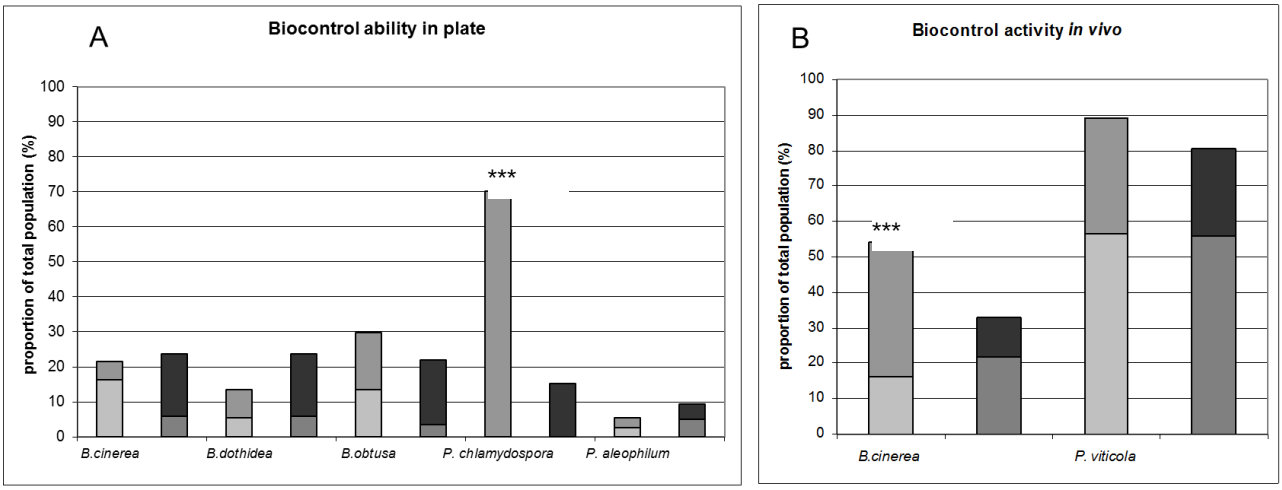


Figure 4. Biocontrol activity of bacterial endophytes from grapevine. Frequency of endophytes from cultivated and wild grapevine (light and dark bars, respectively) inhibiting pathogen growth in Petri dishes (a) and *in vivo* (b). The asterisks (***) represent significant means ($p< 0.005$) in exact binomial test of goodness-of-fit

4.1.2.5. Growth promotion in planta

Some bacterial endophytes tested were highly effective in promoting plant growth, although the test was carried out on non-host conditions. Strains SpVs6 (*Sphingomonas* sp.), PaVv7 (*P. agglomerans*), CuVs2

(*Curtobacterium* sp.), BreVs2 (*Brevundimonas* sp.), BreVs1 (*Brevundimonas* sp.), PsVs4 (*Pseudomonas* sp.), PaVv6 (*P. agglomerans*), PaVs9 (*P. agglomerans*) and PsVs1 (*Pseudomonas psychrotolerans*) significantly increased *A. thaliana*'s biomass. Twenty-seven endophytes significantly increased the number of germinated seeds (Figure 5). Strikingly, among the best strains for enhancement of seed germination, a majority were isolated from domesticated grapevine. Strains from domesticated grapevine were significantly enriched among the best performers as germination enhancer (Chi-square p value is 0.034).

4.1.3. Specific taxa differentiate wild and domesticated grapevine communities

Bacillus: one clade consisted of *Bacillus* sequences clustering with the reference strains *B. subtilis* and *B. amyloliquefaciens*. This clade included 5 strains that were grouped together by phenotype in PCA (Appendix Figure 3). Two out of 7 strains were assigned to this group according to DNA sequence but were not associated to the others by PCA. The same clade included the two strains isolated from domesticated grapevine. A smaller cluster of sequences included the reference strains *B. safensis*, *B. altitudinis*. Another relatively small clade included reference *B. cereus* and *B. pseudomycoides* and other strains showing very high plant growth promotion scores (Appendix Figure 3).

Curtobacterium: the dendrogram shows all strains from wild and domesticated grapevine cluster near the references *C. flaccumfaciens* and *C. herbarum*. Within this clade, strains from wild grapevine appear to form a separate subgroup (Appendix Figure 3).

Pseudomonas: sequences were clustered in clades that included both wild and domesticated grapevine strains. A cluster including many isolates from domesticated grapevine corresponded to the group of *P. poae*, but these strains are not grouped in the same area in the PCA scatter plot (Figure 3). This particular group showed high scores of plant growth promotion, biocontrol and biosensor activation in the quorum sensing bioassays (Appendix Figure 1). Two sequences from domesticated grapevine form a separate clade from the rest of the sequences, these strains are also grouped together by PCA.

Pantoea: most 16S rDNA sequences from genus *Pantoea* were grouped in two main clades. One group included the reference strains *P. agglomerans*, *P. stewartii* and *P. ananatis*. Five wild and two domesticated grapevine endophytes clustered in this group. The other group included the reference strains *P. vagans*. Two strains were close to the reference *P. eucalypti*. We did not observe correspondence of clusters obtained by gene and phenotype analysis.

4.1.4. Remarks

We isolated a variety of endophytic bacteria, comparable to what was observed in previous studies (Goryluk et al. 2009; Zinniel et al. 2002) from surface-disinfected grapevine stems. Roughly, 47% (155 out of 333) of endophytic isolates were considered unique and fully characterised. The taxonomic diversity recovered from wild plants was significantly higher than that in domesticated plants both when the total number and when the number of unique strains were considered. The greater diversity in the endophytic bacterial community of wild grapevines as compared to domesticated grapevines was confirmed by community fingerprinting using B-ARISA (Appendix Figure 2).

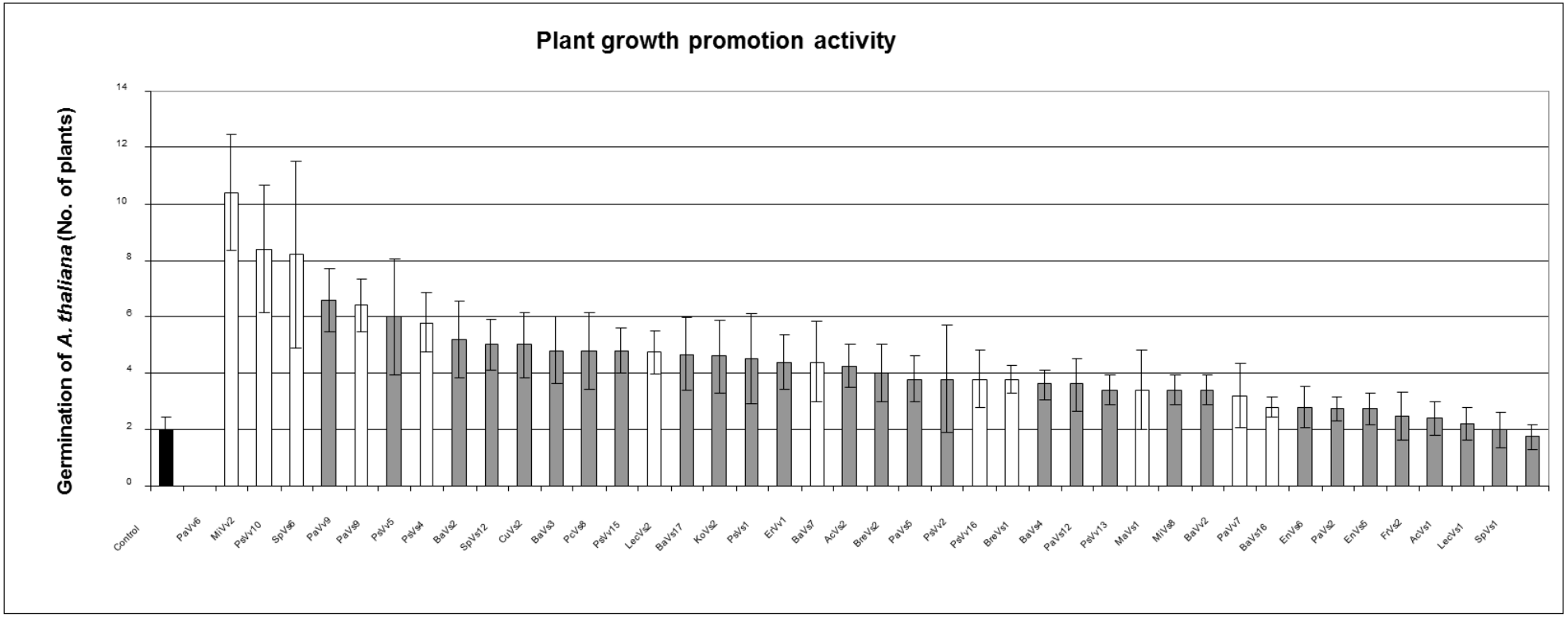


Figure 5. PGP activity of bacterial endophytes from grapevine. Average number of germinated *Arabidopsis thaliana* seeds per plot. White and grey bars represent endophytes from cultivated and wild grapevine, respectively

In the PCA plot made from the whole set of bioprospecting tests (Figure 3a), all strains are grouped in three clusters. All three clusters included strains from both wild and domesticated grapevine plants.

When tested for differences between groups using one-way ANOSIM and one-way NPMANOVA (Hammer et al. 2001), domesticated grapevine endophytes were statistically different from wild grapevine endophytes, in terms of PGP, QS or BiCo tests. The highly significant ($p < 0.001$) multivariate diversity between wild and domesticated grapevine endophytes suggests that these plants, although genetically similar, tend to harbour different microbial communities.

PCA showed that some genera have functions that can be discerned in a host-dependent manner. For example, *Pseudomonas spp*. strains from domesticated and wild grapevines clustered in distinct groups (Figure 3b). *Bacillus spp*. were grouped in two clusters on the PCA (Figure 3c). *Microbacterium* was mostly isolated from wild grapevine and made two separate clusters according to the host. Similar results were obtained for *Pantoea* and *Curtobacterium* (Figures 3d-f).

When we compared phenotype affinity to taxonomy, we found out that taxonomically distant endophytes often clustered together by physiological characteristics, as seen in the multivariate PCA analysis. When comparing phenotypes by PCA, we noted that, within the same genus, endophytes correlated better with those isolated from the same grapevine subspecies than they did with those with similar 16S rDNA sequence,

Results from plant growth promotion and biocontrol tests suggest that these traits are more prevalent among endophytes found in cultivated *V. vinifera* plants, than is their wild relatives. Since propagation by cutting and grafting is likely to promote vertical transmission of endophytes of plant stems, we speculate that the plant's hologenome, rather than the genome alone, selected when domesticating the plant.

Our data show that individuals and groups of endophytes living in the same plant species and belonging to the same genus can be functionally versatile. This observation reinforces the notion that complete microbial community analyses require a polyphasic approach.

4.2. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria

The next step in the study of endophyte molecular ecology was the analysis of genomes in selected bacterial strains isolated from grapevine that performed well in terms of PGP, biocontrol and antibiotic resistance. By sequencing the genomes and taking a look into their structure, we discovered that all the phenotypical properties observed in all tests performed previously (see 4.1) are of course represented at the genomic level. As good as that sounds, the evidence we present next, suggests that careful measures should be taken into account when formulating endophytes in the field, since some pathogenicity attributes are also represented in their genomes.

4.2.1. Sequencing and assembly of endophytic genomes

The characteristics of the genomes sequenced in this study, including length, GC content and amount of transfer and ribosomal RNA genes are summarized in Table 2 and in Appendix Table 2.

Table 2. Characteristics of sequenced genomes from grapevine endophytes and reference genomes

SPECIES	HABITAT/ HOST	LIFE-STYLE	CHROMO-SOME SIZE (BP)	G+C CONTENT (%)	NUMBER OF ORF	REFERENCE
En. cloacae subsp. *cloacae* ATCC 13047	human	Pathogenic humans	5,314,588	54.79	5518	(Ren et al., 2010)
En. asburiae LF7a	human	Pathogenic humans	5,012,130	53.84	4612	(Brenner et al., 1986)
En. aerogenes KCTC 2190	human	Pathogenic humans	5,280,350	54.8	4912	(Shin et al., 2012)
En. sp. 638	poplar	Endophytic	4,518,712	52.98	4240	(Taghavi et al., 2010)
Er. billingiae Eb661	pear tree	Epiphytic	5,100,168	55.2	4587	(Kube et al., 2010)
Er. amylovora ATCC 49946	apple tree	Pathogenic plant	3,805,874	53.5	3483	(Sebaihia et al., 2010)

Er. pyrifoliae Ep1-96	pear tree	Pathogenic plant	4,026,322	53.4	3645	(Kube *et al*., 2010)
Er. pyrifoliae Ejp617	pear tree	Pathogenic plant	3,909,168	53.64	3873	(Park et al., 2011)
P. agglomerans 299R	pear tree	Endophytic	4,581,483	54.29	4194	(Remus-Emsermann et al., 2013)
P. ananatis LMG 20103	eucalyptus tree	Pathogenic plant	4,690,000	53.69	4241	(De Maayer et al., 2010)
P. ananatis PA13	rice plants	Pathogenic plant	4,586,378	53.66	4372	(Choi et al., 2012)
P. vagans C9-1	apple tree	Epiphytic	4,025,000	55.09	4619	(Smits et al., 2010)
En. Ludwigii. EnVs6	grapevine plants	Endophytic	5,220,112	54.62	4809	This study
En. ludwigii EnVs2	grapevine plants	Endophytic	5,067,900	53.98	4649	This study
En. ludwigii LecVs2	grapevine plants	Endophytic	5,285,925	54.59	4886	This study
Erwinia sp. ErVv1	grapevine plants	Endophytic	4,719,019	54.6	4207	This study
P. vagans PaVv1	grapevine plants	Endophytic	4,850,774	55.14	4453	This study
P. vagans PaVv7	grapevine plants	Endophytic	4,879,255	55.22	4470	This study
P. vagans PaVv9	grapevine plants	Endophytic	9,754,510	54.65	9466	This study

The sequencing of the 16S rDNA confirmed that the test strains in the genus *Enterobacter* are closely related to *En. ludwigii*; strain ErVv1 in the genus *Erwinia* is related to *Er. amylovora* and *Er. tasmaniensis*, while the *Pantoea* test strains are closely related to *P. vagans* (Appendix Figure 4).The sequencing data is deposited in the EMBL-EBI repository (https://www.ebi.ac.uk/ena) with accession numbers PRJEB8251 (EnVs6); PRJEB8253 (EnVs2); PRJEB8254 (LecVs2); PRJEB8255 (PaVv1); PRJEB8258 (PaVv7); PRJEB8259 (PaVv9) and PRJEB8284 (ErVv1).

To optimize the assembly procedure, we compared the performance of two different pipelines (A5 and SOAPdenovo) through the online tool QUAST (Appendix Table 3). The realignments of the reads showed consistently good results with high realignment percentages and the distribution of the insert sizes within the limits of the suggested values from the sequencing provider (~600/800bp). The Log average Probability (LAP) scores (Ghodsi et al., 2013) from the different assemblies for the same organism were also very close, showing that both pipelines produced comparable results. We used these scores and the N50 values to choose which assembly pipeline to use. For strain LecVs2, pipeline A5 did not produce any results. Thus we used the Velvet pipeline (Zerbino and Birney, 2008) to assemble the genome and compared it with the SOAPdenovo assember. Quality analysis showed that differences between the two pipelines were minimal as presented in Appendix Table 3. Scripts to build up the assemblies are deposited and publicly available at the Github online public repository (at the address https://github.com/pochotustra/genomics_endophytes.git).

4.2.2. Comparison of genome structure in test and reference strains shows lifestyles and chromosome arrangement are linked

A visual inspection of the circular alignment of genomes in the genus *Enterobacter* (Figure 6A-B) highlights that two of the test genomes (EnVs2 and LecVs2) are similar to the alignment reference genome of endophytic strain 638, while the test strain EnVs6 is similar to the reference genomes of human pathogens (ATCC 13047, LF7a and KCTC 2190). The region between 1-800 kbp is well conserved in all three test isolates (EnVs6, EnVs2 and LecVs2). The rest of the chromosome is more variable, with regions where the identity between EnVs2 and LecVs2 reaches up to a 70% while the test strain EnVs6 aligns better with the genomes of pathogens (ATCC 13047, LF7a and KCTC 2190). A positive to negative GC skew at position 1 corresponds to the origin of replication and the switch at 2200 kbp may account for the replication terminus. Regions at 2800 - 2880kbp and 3120 - 3200 kbp have higher content of GC. We also found that more than 10 regions in the genome of strain 638 and of the test strains (EnVs2 and LecVs2) are absent in the pathogenic reference strains and in one test strain (ATCC 13047, LF7a, KCTC 2190 and EnVs6). These missing regions contain several enzymes involved in nitrogen assimilation including genes coding for a 2,3, 4,5-tetrahydropyridine-2,6-dicarboxylate *N*-succinyltransferase, a uridylyltransferase and several enzymes for the synthesis of the core carbohydrate 3-deoxy-D-manno-octulosonic acid (KDO).

In the genus *Erwinia* a positive to negative GC skew at position 1 in Figure 6C corresponds to the origin of replication and a positive to negative GC switch at approximately 1920 kbp corresponds to the replication terminus in all the genomes compared. We found seven positions with higher GC contents in regions 440-480 kbp, 600-960 kbp, 1280-1320 kbp, 1640-1720 kbp, 2200-2640 kbp, 2880-2920 kbp and 3340-3400 kbp. These regions contain several virulence factors including *vgrG*, *virB* and several *imp*, all genes related to the type VI secretion system (Filloux et al., 2008). Also, several genes of the biotin biosynthesis pathway and a few DNA repair genes like the methylated DNA protein cysteine methyltransferase are found in these regions, according to the location on the genome of the alignment reference genome ATCC 49946.

A qualitative interpretation of the plots suggests that the genome of strain ErVv1 is similar to that of the alignment reference genome ATCC 49946 (a plant pathogen) and they share several regions that are not present in the other genomes compared. Conversely, the reference strains Ep1-96, Eb661 and Ejp617 are more similar to each other. The more evident gaps highlighting the missing regions are visible at positions

40-60 kbp, 440-530 kbp, 2400-2440 kbp, 2920-2960 kbp, 3120-3160 kbp, 3220-3240 kbp and 3360-3400 kbp (Figure 6C).

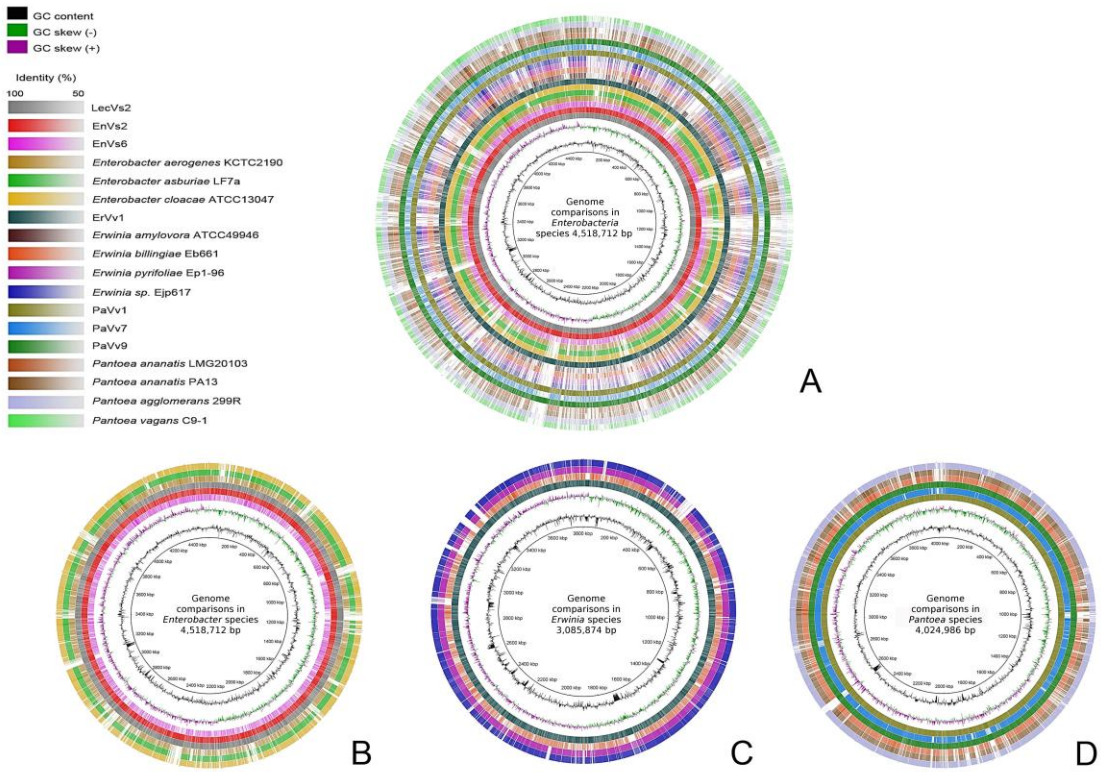


Figure 6. Whole-genome comparisons in three genera of Enteobacteria. The color intensity in each ring represents the BLAST match identity. **A)** whole-genome comparison of all the strains considered in this work. **B)** whole-genome comparisons in *Enterobacter***,** from outer to inner ring: *En. cloacae* subsp. *cloacae* ATCC 13047, *En. asburiae* LF7a, *En. aerogenes* KCTC 2190, *En. ludwigii* LecVs2, *En. ludwigii* EnVs2, *En. ludwigii* EnVs6; reference genome: *Enterobacter* sp. 638**. C)** whole-genome comparisons in Erwinia, from outer to inner ring: *Er.* ErVv1, *Er. pyrifoliae* Ejp617, *Er. pyrifoliae* Ep1-96, *Er. billingiae* Eb661; reference genome: *Er. amylovora* ATCC 49946. **D)** whole-genome comparisons in *Pantoea*, from outer to inner ring: *P. agglomerans* 299R, *P. ananatis* PA13, *P. ananatis* LMG 20103, *P. vagans* PaVv9, *P. vagans* PaVv7, *P. vagans* PaVv1; reference genome: *P vagans* C9-1.

In Figure 6D test and reference strains of the *Pantoea* group are compared. The test strains PaVv1 and PaVv9 are similar to the C9-1 strain (a biocontrol agent) used as alignment reference genome, while the test strain PaVv7 is very similar to the genomes of pathogens PA13 and LMG20103. We found regions with higher GC content at positions 960-1120 kbp, 1800-2200 kbp, 2560-2600 kbp and 2840-2920 kbp. The GC switches at 1440 kbp (positive to negative) and at 3480 kbp (negative to positive) may represent the origin of replication and replication terminus respectively (Figure 6D). The sequence identity between the test strains (PaVv1, PaVv7 and PaVv9) and the alignment reference genome C9-1 is high throughout the alignment and can reach up to 100% identity. This is more evident in the regions between 200 -1700 kbp and 2400 -4000. Several regions shared between the alignment reference genome and the test strains are instead absent in the genomes of pathogens PA13 and LMG20103. The genes differentially present in these regions are

involved in cysteine metabolism, isoprenoid biosynthesis and in the transport and metabolism of D-glucarate (ascorbic acid biosynthesis).

Through synteny plots, preservation of chromosomal organization was checked. Synteny is evidenced by the sharpness of the line in the plot. Several regions in the genome of test strains (EnVs2 and LecVs2) are synthenic with the genome of the alignment reference genome of endophytic strain 638. For strain EnVs6 however this synteny is less conspicuous as it is made evident in the alignment plot (Appendix Figure 5). Regions between 240 - 280 kbp and between 4040 - 4080 kbp in the test strains (EnVs2, EnVs6 and LecVs2) show an identity higher than 70% to the alignment reference genome of strain 638. In these regions we located genes related to the thiamin biosynthesis (the thiamin phosphate pyrophosporilase and the thiamin biosynthesis gene *thiC*) and several genes of the *yjb* operon that regulate the synthesis of a stress-induced exopolysaccharide in *E. coli* (Ionescu et al., 2008).

Synteny plots constructed for the test strain genomes in the taxon *Erwinia* show a high degree of homology with the alignment reference genome. The number of insertions or deletions is low, which is confirmed by the continuity in the plots. For strain ErVv1 the number of discontinuities is higher (6 regions of the genome being separated by indels; Appendix Figure 5).

Synteny plots made for *Pantoea* show homology of test strains to the alignment reference genome C9-1. We located two exceptional re-arrangements (indels) in strains PaVv9 and PaVv1 that are not present in strain PaVv7, Appendix Figure 5)

4.2.3. Core and accessory genomes from endophytes are enriched in virulence factors

Core and accessory genomes were calculated and compared in order to find commonalities and differences between enterobacterial genomes of strains sequenced in this study.

4.2.3.1. Genus Enterobacter

Our analysis shows that the core genome in this taxon comprises 2468 orthologous genes that correspond to 53% of the pangenome (Figure 7A). The abundance of genes in each category (Appendix Table 4 contains a list of categories used) was similar in all genomes (Appendix Table 5).

Cell signaling and two-component system genes are present in all of the strains analyzed. In these systems, transcriptional regulators belonging to the *lysR* and *gntR* families (Fujita and Fujita, 1987; Maddocks and Oyston, 2008) are the most important. We found a conserved set of ca. 40 genes devoted to the synthesis of cell wall and capsule that are shared among all *Enterobacter* genomes analyzed. Among these, we highlight the presence of *rlpB, rcsF*, *lptA* , *lptC* and the organic solvent tolerance protein *lptD* a set of genes involved in LPS biosynthesis (McCandlish and Silhavy, 2007) and the oxidoreductases *mviM* and *mviN*, necessary for murein synthesis (Inoue et al., 2008). We found an average of 51.8 flagellar genes across the *Enterobacter* genomes with a maximum in the reference pathogenic strain ATCC 13047. A set of genes that belong to the *che* operon for chemotaxis signaling and to the twitching motility apparatus are also present in the core genome. The number of genes related to pathogenicity mechanisms was slightly higher in the genomes of the test strains LecVs2 and EnVs6 and in the genome of the reference genome ATCC 13047. Among these genes we emphasize the presence of a virulence sensor related to the *bvgS* sensor kinase and of arnC

belonging to the polymyxin resistance group. We report the presence of a type 1 secretion system agglutinin of the RTX family (Linhartová et al., 2010) and members of a tripartite multidrug resistance system. Some of the genes in this category are also related to the type II secretion systems and genes for the biogenesis of type IV pillus.

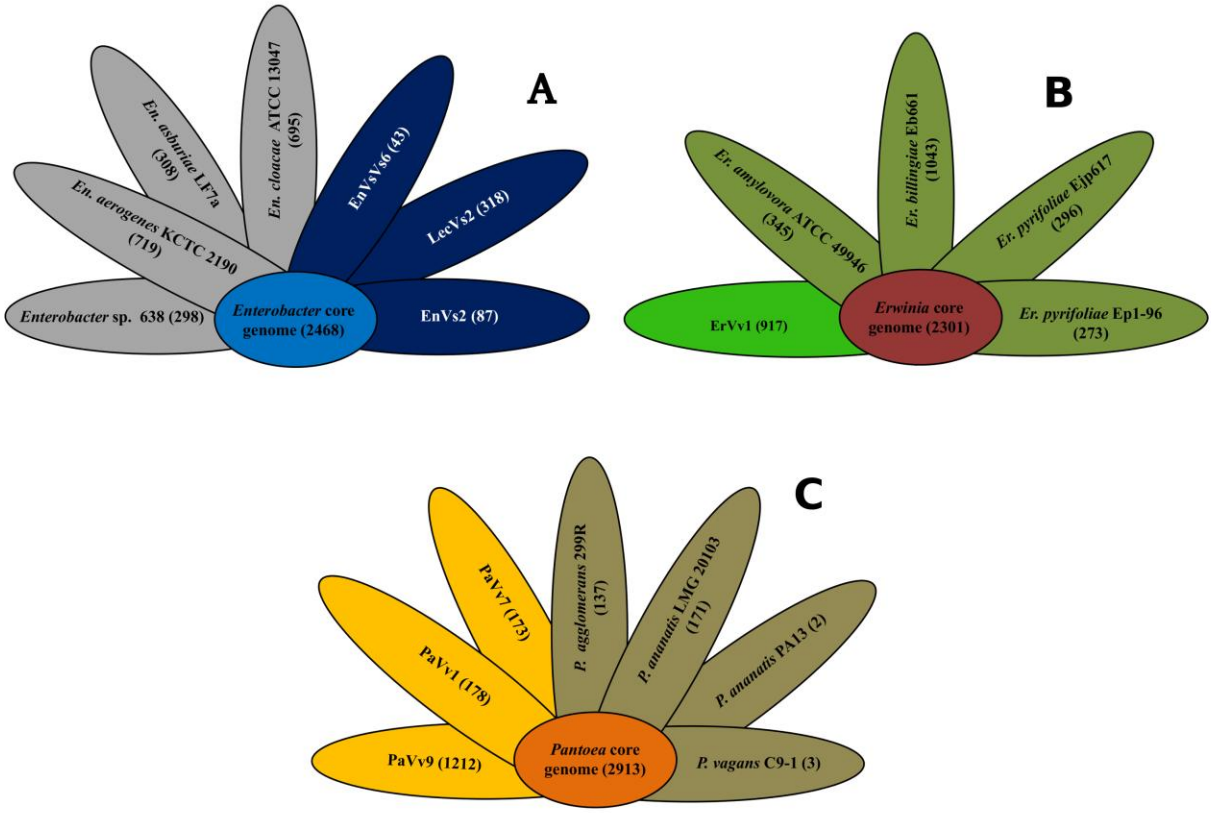


Figure 7. Core and accessory genome size in endophytic and non endophytic bacteria. **A)** Core (sky blue) and accessory genome size of reference (grey petals) and test (deep blue) strains of the genus *Enterobacter*. **B)** Core (terracotta) and accessory genomes of reference (olive petals) and test (bright green petals) strains of the genus *Erwinia*. **C)** Core (orange) and accessory genomes of reference (brown petals) and test (yellow petals) strains of the genus *Pantoea*. Numbers indicate genes counts**.**

The categories with the lowest number of genes were phages and quorum sensing in which the highest number of genes among the test strains was in the genome of EnVs6 (Appendix Table 5). Several CDS for phage capsid and phage associated enzymes (terminases and integrases) are shared among all strains. In the quorum sensing category, the *sdiA* gene of the orphan quorum sensing communication circuit from *E. coli* (Kanamaru et al., 2000) was detected. Also a *N*-acyl homoserine lactone synthase was detected in all genomes. No AI-2-dependent quorum sensing systems were detected in the test strains. We found no variation in the content of siderophore related genes (average 52,14 genes). However the test strains (EnVs6, EnVs2 and LecVs2) contain a higher number of genes for this category as compared to all the reference strains (Appendix Table 5). We found genes for enterobactin synthesis and some of the genes for

hemin metabolism. Also, exo- and endoenzymes are present in the core genome of the *Enterobacter* including the phospholipase A and two hemolysins.

Venn diagrams show genes exclusively shared between test strains and either of the reference genomes. In strains EnVs6, EnVs2 and LecVs2, the number of genes shared exclusively with the endophytic reference strain 638 was 149, 127 and 150, respectively. The number of genes shared exclusively with the pathogenic reference strain (ATCC 13047) was 254, 257 and 256, respectively (Figure 8A).

4.2.3.2. Genus Erwinia

In the genus *Erwinia*, the core genome is constituted by 2301 gene clusters that correspond to the 60% of the pangenome (Figure 7B). The core genome of the genus *Erwinia* defined in this study is populated with cell signaling functions including two component systems sensitive to nitrates, copper and osmolarity (Appendix Table 6). We located also two transcriptional factors belonging to the Rrf2 family involved in the metabolism of cysteine (Shepard et al., 2011) and a cyclic AMP regulator of the *crp*/*fnr* family (Shimada et al., 2011). Among the most important gene functions found for cell wall is the regulator in colanic acid synthesis that confers a mucoid phenotype in other taxa. The flagellar machinery is present in all the strains analyzed. As pathogenicity mechanisms we detected multidrug efflux transporters as well as heavy metal detoxification genes including the *arcB* gene (Iuchi et al., 1990) and the cation efflux pump *fieF* (Munkelt et al., 2004); the genus also bears well known antibiotic inactivating proteins, including the ampG beta lactam activation protein (Lindquist et al., 1993) and both *mdtK* and the multidrug resistance gene *emrD* (Lomovskaya and Lewis, 1992). Several phage elements and the sdiA gene from the quorum sensing circuit along with a *N*-acyl homoserine lactone synthase and a homoserine lactone transporter are conserved in all species analyzed. The core genome contains also general mechanisms for iron acquisition that comprise 36 different functions. In the exoenzyme category we found two cellulose synthesis genes, one phospholipase A related gene and a cryptic hemolysin regulator.

Venn diagrams in Figure 8B show genes exclusively shared between test strains either the endophytic reference genomes or the pathogenic reference genomes. We found that strain ErVv1 shares 133 genes exclusively with the apple tree pathogen ATCC 49946 and shares 619 genes exclusively with the epiphyte Eb661

4.2.3.3. Genus Pantoea

In the genus *Pantoea* we found 2913 orthologous families. The core genome corresponds to 60% of the pangenome (Figure 7C).

The cell signaling mechanisms in the core genome of *Pantoea* include the diguanylate cyclase mediator of biofilm formation, along with other biofilm-related genes like *rcsB* (Shiba et al., 2006), and one ribose metabolizing gene rpiR (Sorensen and Hove-Jensen, 1996). In this category two distinct groups: one made of 299R, LMG 20103, PaVv1 and PaVv7 and a second group made of PaVv9, C9-1 and PA13 are well defined (Appendix Table 7) The cell wall and capsule functions form the same two groups, in which the second group has a greater number of genes in the core genome (Appendix Table 7). The functions for these genes are linked to the synthesis of the exopolysaccharide substance amylovoran and of lipopolysaccharide

modifications. The presence of a putative cellulose synthesis gene is consistent in all the species. As for pathogenicity mechanisms, we found that the number of genes in strain PaVv9 is the highest among the test strains (282 genes) and similar to the number found in the reference plant pathogen PA13 (290 genes). Also a high number of genes for flagellum assembly (139 genes), exo- and endo-enzymes (117 genes), capsule (141 genes) and phages (37 genes) are present in strain PaVv9 as compared to the other test strains and to some of the reference genomes (Appendix Table 7). Inside the core genome, we located the *shlA* gene and several other putative adhesines. We also found the pilN gene of pilus biogenesis (Sakai and Komano, 2000) along with the two partner secretion system tpsA/B that corresponds to a conserved virulence factor for host adhesion and toxicity in bacterial cells (Rahman and van Ulsen, 2013). Finally we provide evidence for the presence of the phenazine synthesis gene phzF (Parsons et al., 2004).Some of the virulence-related gene products are transmembrane transporters including members of the major facilitator superfamily (MFS) and members of the Tol-Pal system involved in membrane integrity and phage acquisition. The core genome of the *Pantoea* is poor in genes related to iron acquisition although it contains important elements belonging to the ABC transporters and to the TonB-dependent pathway. Also in the core genome of the *Pantoea* species analyzed in our study, several exo- and endo-enzymes were detected. Particularly, we were able to track the bcs genes involved in cellulose metabolism. Several regulatory proteases make part of the core genome of the genus. We found *hsl*, *clp*, *lon* and *fts* present in all the genomes analyzed. We also detected a gene of the protein lipase 1LIP-1 and the collagenase *yhbU* that has been implicated in virulence of several animal pathogens. We report also the finding of a *luxI*-type coding gene, similar to the pagI synthase and a transcriptional activator belonging to the LuxR family that is related to the N-3-oxohexanoyl-L-homoserine lactone quorum-sensing transcriptional activator of *P. agglomerans* (Appendix Table 7).

Venn diagrams show that strain PaVv7 shares 549 genes exclusively with the symbiont C9-1 while test strains PaVv1 and PaVv9 share 562 and 653 respectively. One-hundred and sixteen genes are shared exclusively between the rice pathogen PA13 and the test strain PaVv7 while test strains PaVv1 and PaVv9 share 102 and 166 genes respectively (Figure 8C).

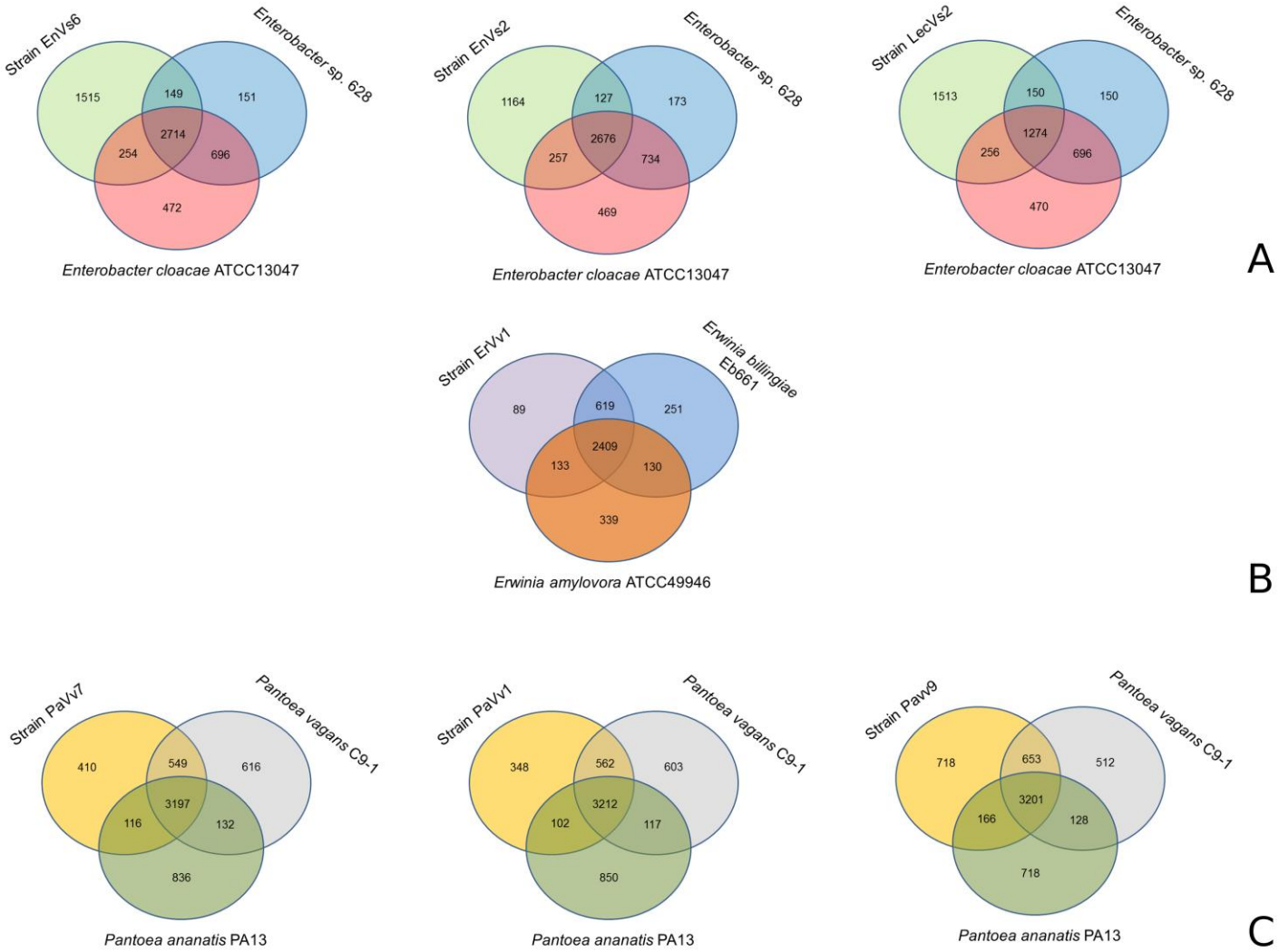


Figure 8. Venn diagrams showing shared and unique genes in *Enterobacteriaceae.* All comparisons are made between the test strains (left circle) with an epiphyte or endophyte (right circle) and a pathogen (lower circle).

A) Comparison of strains EnVs6, EnVs2 and LecVs2 with the endophyte *Enterobacter.* sp. 638 and the human pathogen *En. cloacae* ATCC 13047.

B) Comparison of strain ErVv1 with the epiphyte *Er. billingiae* strain Eb661 and the plant pathogen *Er. amylovora* ATCC 49946.

C) Comparison of strains PaVv7, PaVv1 and PaVv9 with the epiphyte *P. vagans* C9-1 and the plant pathogen *P. ananatis* PA13. Numbers inside the circles, indicate the genes shared among genomes.

4.2.4. Unique gene functions in the different lifestyles

In each genus, a list of shared and unique features of virulence-related genes was drawn (Table 3). The genomes of both reference and test strains in all genomes contain most of the main functions for cell wall synthesis and flagellum synthesis which are basic for the niche occupation. They are characterized however by minor differences in which some genes are unique for test strains or for reference strains, respectively.

In the genus *Enterobacter* genes involved in the modification of lipopolysaccharide as well as genes related to the glicocalix synthesis and modification are lacking in the test strains. Genes related to the use of sialic acid are present only in the test and absent in the reference genomes. We also found that genes involved in the metabolism of hemin and synthesis of aerobactin are absent in the test and present in the reference genomes. We detected variable content of flagellar genes since test strains possess one component of the flagellum apparatus uniquely present and lack one gene of the *che* group.

In the genus *Erwinia* we highlight four genes that are absent in the genomes of the reference strains with unrelated functions: antibiotic resistance, detoxification of xenobiotics and stress response. In the genus *Pantoea* we stress the presence of genes that are unique only in the test strains. These genes are necessary for plasmid stability and for recognition and resistance to exogenous substances. On the other hand, the test strains lack genes for modification of the KDO in the LPS and one sequence coding for structural components of a phage, that are only present in the reference strains (Table 3).

Table 3. Unique functions in the genomes of Enterobacteria. (*) = strains include endophytic reference genomes *Enterobacter* sp. 638 and *Pantoea vagans* C9-1 and all 7 test strains. Presence (+) and absence (-) of the genes is denoted.

GENUS	CATEGORY	GENE IDENTIFIER	FUNCTION	TEST STRAIN	REFERENCE STRAIN	ENDOPHYTIC STRAINS*
Enterobacter	Chemotaxis	operon kps	sialic acid ABC transporters	+	-	-
		rgpF	alpha rhamnosyl transferase	+	-	-
		nanTRC	synthesis and modification of sialic acid	-	+	-
		operon wca	synthesis of colanic acid	-	+	-
		yjbG	glycocalix modification	-	+	-
		yjbF	glycocalix modification	-	+	-
		waaL	O-antigen ligase	-	+	-
		iutA	aerobactin transporter	-	+	-
		hypothetical	hemin metabolism	-	+	-
		flaA	flagellin assembly	+	-	-
		che group	methyl-accepting proteins	-	+	-
	Phages	phage gene	phage baseplate hub necessary for host recognition	+	-	-
	Regulation and cell signalling	*ibrA/B*	immunoglobuling inactivating protein	-	+	-
		higA	toxin - antitoxin	+	-	-
		parD	toxin - antitoxin	+	-	-
		vapB	toxin - antitoxin	+	-	-
		gbuR	arginine dehydrogenase transcriptional regulator	+	-	-

		evgA	two component system kinase	+	-	-
		operon lsrB	AI-2 metabolism	-	+	-
		lsrC	AI-2 metabolism	-	+	-
		lsrD	AI-2 metabolism	-	+	-
		lsrG	AI-2 metabolism	-	+	-
		lsrF	AI-2 metabolism	-	+	-
		lsrK	AI-2 metabolism	-	+	-
		ccpA	carbon control protein A	-	+	-
		tdcA	threonine catabolic transcriptional activator	-	+	-
	Virulence and disease	operon *pga*	poly-β-1,6-N-acetyl-D-glucosamine (PGA) synthesis	-	+	-
		copS	the copper sensor protein	+	-	-
		espC	enterotoxin	+	-	-
		mtrF	multidrug efflux pump	+	-	-
		yidP	transcriptional regulator	-	+	-
		yidE	mediator of hyper adherence	-	+	-
		marR	multiple-antibiotic resistance	-	+	-
		marB	multiple-antibiotic resistance	-	+	-
Erwinia	Cell wall and capsule	hypothetical	acyl hydratase	+	-	-
		ycfS	L,D-transpeptidase	-	-	+
		ycfL	protein: an outer membrane lipoprotein that is part of a salvage cluster	-	-	+

		ycfP	protein: probably an esterase that is part of a salvage cluster	-	-	+
		mpaA	Gamma-D-Glutamyl-meso-Diaminopimelate Amidase	-	-	+
		mltB	Membrane-bound lytic murein transglycosylase B	-	-	+
		msrA	Peptide methionine sulfoxide reductase	-	-	+
		cobU	Adenosylcobinamide kinase	-	-	+
		ytfR	Putative sugar ABC transport system, ATP-binding protein	-	-	+
		ytfQ	Putative sugar ABC transport system	-	-	+
		yjfF	Putative sugar ABC transport system, permease protein	-	-	+
		ytfT	Putative sugar ABC transport system, permease protein	-	-	+
		ytfP	Gamma-glutamylcyclotransferase family protein	-	-	+
		MFS	Major facilitator superfamily (MFS) transporter	-	-	+
		hypothetical	probable Rhodanese-related sulfurtransferase	-	-	+

		pmrC	release factor (RF) methylating enzyme	+	-	-
	Iron metabolism	hypothetical	probable Ferric reductase	-	-	+
		pntA	NAD(P) transhydrogenase alpha subunit	-	-	+
		pntA	NAD(P) transhydrogenase subunit beta	-	-	+
	Phages	*mnmC*	5-methylaminomethyl-2-thiouridine-forming enzyme	-	-	+
	Regulation and cells signaling	GTH	glutathione S- transferase subgroup 2	+	-	-
		cstA	carbon starvation protein A	+	-	-
		yfcF homolog	Probable glutathione S-transferase	-	-	+
		hypothetical	Uncharacterized glutathione S-transferase-like protein	-	-	+
		NQO	NADPH:quinone oxidoreductase 2	-	-	+
		qorR	Redox-sensing transcriptional regulator	-	-	+
		hmp	Flavohemoprotein (Nitric oxide dioxygenase)	-	-	+
		soxS	Regulatory protein	-	-	+
		ytmO	Bacterial luciferase family protein	-	-	+

		hypothetical	ABC-type nitrate/sulfonate/bicarbonate transport systems	-	-	+
		hypothetical	probable dibenzothiophene desulfurization enzyme	-	-	+
Pantoea	Cell wall and capsule	*rfbF*	dTDP rhamnosyl transferase	-	+	-
		waaX	core sugar synthesis in the LPS	-	+	-
	Iron metabolism	*fhuC*	Ferrichrome transport ATP-binding protein	-	-	+
	Phages	mb-GpV	baseplate assembly protein V	-	+	-
	Regulation and cell signaling	*relB-relE*	toxin - antitoxin	+	-	-
		yedV/yedW	two component system	+	-	-
	Virulence and disease	*cusC*	copper sensitive and detoxifying efflux pump	+	-	-
		cusF	copper sensitive and detoxifying efflux pump	+	-	-
		cusR	copper sensitive and detoxifying efflux pump	+	-	-
		operon pmr	Antibiotic resistance	+	-	-
		czc	detoxifying functions for Cd2+, Zn2+ and Co2+	+	-	

4.2.5. Other virulence determinants: CRISPRs and phage sequences

We identified only two CRISPR systems, located in the genome of strain EnVs6 (Table 4). The systems consist of two sets of repeats of approximately 20 nucleotides each. The first set consisting of only three spacers and four repeats and the second system with seven repeats and six spacers. No CRISPRs were identified in the genomes of strains LecVs2, EnVs2, ErVv1, PaVv1, PaVv7 orPaVv9.

Table 4. CRISPRs found in the genome of strain EnVs6.

RANGE	POSITION	REPEAT	SPACER
CRISPR 1 Range: 556393 - 556547	556393	CGCCATTCATGGCGACCTT	ATTAATGGCCGCACCCTGCCCCG
	556435	CGCCATTCATGGCGACCTT	CTTCTTACACGCACCCAACNTAATCCGTAGGGT
	556487	CGCCATTCATGGCGACCTT	ATTAATGGCCGCACCCTGCCCCG
	556529	CGCCATTCATGGCGACCTT	
CRISPR 2 Range: 2949702 - 2949985	2949702	ATGTTCACTGTAATCAGTAAA	ACTTGATGACTTTTCTTTCTCAACGCCTA
	2949752	ACATTCACTGTAATCAGTGAA	AACCTTGTGCTCATCATAGACAA
	2949796	AGATTCACTGTAGTCAGTAGA	TGGTTATCGCGCTTCGATTTA
	2949838	ACGTTCACTGTAATCAGTAAG	TCAAATCTGCAACTTCGACAGA
	2949881	ACATTCACTGTAATCAGTAAA	AGTTATGACCCGGAGAAGA
	2949921	ACGTTCACTGTAATCAGCAAA	GCTTGTTTAGTACTTTGATACGA
	2949965	GCGTTCACTGTAATCAGTAAA	

With RAST we detected several phage-related proteins including tail and sheath proteins as well as DNA modifying proteins of phage origin (Appendix Tables 5-7), We used PHAST to confirm the presence of phages that we were able to identify with RAST. With PHAST we were able to identify 12 intact phages in the genomes of the test strains EnVs6, LecVs2, EnVs2, PaVv9, PaVv1, and PaVv7 (Appendix Table 8). The phage sequences belong to different types of phages, and are identical to those present in Enterobacter genomes deposited in NCBI. The phages do not seem related to one another.

We detected some of the components for major secretion system in all test strains (Appendix Table 9). In all the test strains we detected some of the flagellar components plus the protein clpV. In Enterobacter, all test strains contain the hemaglutinin/adhesin gene *shlA* (Poole and Braun, 1988). In *Erwinia* we did not find any common secretion system gene among the test and reference strains and in *Pantoea* test strains we found the gene coding for the protein Sfa3 of the type VI secretion system apparatus in *Pseudomonas aeruginosa* (Sana et al., 2013).

4.2.6 Remarks

Our study clearly shows how endophytic test genomes are not only similar to reference endophytes but also share several characteristics with pathogenic reference genomes (Figure 6 and apendix Figure 5). At the structural level, we have demonstrated a high degree of synteny between endophytic test and reference strains (Appendix Figure 5) but also a high identity percentage at the genome level between endophytes and symbionts with other lifestyles (Figure 6). We hypothesize that these traits reflect the potential of bacterial endophytes to express virulence when associated with their hosts. In *Pantoea* for example, the similarities at the structural level between endophytes and plant pathogens are linked to the absence of regions that contain key enzymes for aminoacid biosynthesis and vitamin production. In *Enterobacter* it is the absence of genes for central metabolism and of some, but not all, genes for exopolysaccharide modification (that are otherwise present in the endophytic reference strain 638) and in *Erwinia* the similarities between test and pathogenic reference genomes are based on the synteny between the genomes of the test strain ErVv1 and the reference strain ATCC 49946 and on the presence of virulence genes from pathogenicity islands (a trademark genome arrangement of virulent phenotypes (Appendix Tables 5-7). These genomic characteristics might be a sign of lifestyle switching (from endophytes to pathogens and vice versa) when conditions are optimal for such phenomenon. For example, strains that have no enzymes for ascorbic acid production might be prone to take it from the environment, thus scavenging the substrates from the host's cells (Abu Kwaik and Bumann, 2013). Also, the presence of pathogenicity island components might suggest the use of such modules to achieve colonization using a pathogen-like strategy.

Our comparisons also show that, endophytes, epiphytes and pathogens share a wide number of virulence-related genes. We found that core genomes are densely populated with virulence factors. These genes are present in each of the genomes in this study (Figures 6-7) and they mark a baseline for the existence of a **core virulence genome**. Furthermore, the virulence-related genes found in the core genomes are conserved within each genus, regardless of pathogenic or endophytic lifestyles (Appendix Tables 5-7). For example, the core genome of the genus *Pantoea* is characterized by the largest number of virulence-related genes and genes existing in this taxon are not present in the other groups analyzed. This supports our hypothesis that differences between endophytes and pathogens do not exist *per se*, and demonstrates that the similarities between these two groups are set above the species level. Previous research found that mutations in some genes can attenuate the virulent phenotype of some strains (Matthysse et al., 1995) and it is then conceivable that such mutations or may produce an endophytic phenotype as an attenuated virulence state. We extend this rationale to other genes of the core genomes, like those coding for the phospholipase A, a well-known virulence determinant that in human pathogens is activated upon antagonistic antimicrobial activity (Istivan and Coloe, 2006); the *clpV* gene that has been linked to the proper folding of effector proteins in pathogenic strains (Schlieker et al., 2005;Filloux et al., 2008) and those virulence-related genes that are key in regulatory networks in other taxa, including the orphan QS gene s*diA* (Kanamaru et al., 2000) found in all the strains analyzed. Although our methods for calculating virulence factors may have introduced a bias, the genes that we have found in this study reflect existing and previously reported functions. Together with the aformentioned findings, our observations are consistent with the emerging idea of pathobiome and 'balanced antagonism', by which host-adapted bacteria can play different roles, depending on their relations with the environment.

Complementary to the core genome, the accessory genome reveals differences in niche specialization. Variation in accessory genome size might be related to the endophytic lifestyle, since genome reduction is reported in symbiotic microorganisms (McCutcheon and Moran, 2012). Strain EnVs6 for example, displays the smallest accessory genome in the *Enterobacter* set of strains (Figure 7A), fitting the concept that endophytes have reduced genomes as compared to pathogens. In contrast, the accessory genome of the test strain ErVv1 is large, paralleling the one of the epiphyte strain Eb661 (Figure 7B). Also, genome structure in the test strains resembles that of endophytic reference genomes (Figure 6) in the number of genes that compose the accessory genome. For example, in Figure 7A the accessory genomes of strain EnVs2 and EnVs6 contain lower gene numbers, similar to what happens in the endophytic reference genomes *Enterobacter* sp. 638. The same happens in strain ErVv1, which contains about the same number of genes as the endophytic *E. billingiae* Eb661 (Figure 7B) and in strains PaVv1 and PaVv7 that are similar to the endophytic reference *P. agglomerans* 299R (Figure 7C). Niche specialization might be related to the abundance of gene functions like the toxin-antitoxin systems (Appendix Tables 5-7) that are thought to be more abundant in free living engaging bacteria and fewer in other types of lifestyle (Pandey and Gerdes, 2005) and the ABC-type polar amino acid transport system for opine translocation (Moore et al., 1997) that in other taxa is important for selection of bacterial pathogenic subpopulations.

We also present a set of genes that could be found only in the endophytic genomes, either test or reference strains (Table 3). The nature of these genes is quite diverse and span from functions related to the normal modification of the cell surface to the catabolism of by-products of the S-adenosyl cystein pathway. However, it is yet to be discovered how such modifications might be involved in the mechanisms of colonization or if they are used for beneficial associations or for pathogenicity.

We propose that some of the traits found provide clues on how bacteria with an endophytic lifestyle maintain a symbiotic relation with its host as they are present sometimes only in endophytes and moreover in grape endophytes (test strains). This is the case of endophytic phage sequences (Appendix Table 8) which are not rare in the endophytic genomes (Ozer et al., 2014). We speculate that a link between phage sequences and the attenuation of virulence in endophytes might exist given the widespread appearance of such sequences only in the genomes of test strains. Moreover, a genome analysis of the endophytic test strains revealed that the core endophytic genome (i.e. the collection of orthologous genes present only in the endophytic test strains of our set) contains only 536 gene families (the endophytic core genome of our test strains) and is populated with functions related to vitamin synthesis and to cell signalling as well as virulence (data not shown). This suggestins that endophytic only a limited part of the genome of endophytes is dedicated to sych associations and that virulence is a leading trait in that core genome.

Summing up these observations, we suggest that endophytes conserve properties of different lifestyles, including pathogenic traits. This is reflected in the structural organization of the genomes (Figure 6) and in the overlapping functions between the test strains and the genomes of plant or animal pathogens, epiphytes or endophytes (Figure 7). We propose that endophytic and pathogenic lifestyles are composed of a base core virulence genome that might be used and expressed differentially, as has been shown for other taxa (Meysman et al., 2013). Lifestyle in pathogens or endophytes might be the outcome of a complex,

multifactorial interaction. Our conclusions are consistent with the hypothesis that relationships between environment, host and microorganism(s) contribute to shape the environmental role of microorganisms in this symbiosis, independent of their phylogenetical relatedness. Our research is to the best of our knowledge a pioneer in two regards. First, we are showing similarities between sequenced genomes of endophytic strains from grapevine while also emphasizing on the differences that our endophytic test strains present when compared with organisms spanning other lifestyles. Secondly, we are using comparative genomics to establish a link between the genome content and genome organization of endophytic (beneficial) organisms with niche occupation, by highlighting the role of specific characteristics of the genome, that lead to different degrees of specialization.

4.3. A phloem-feeding insect transfers bacterial endophytic communities between grapevine plants

After having confirmed beneficial properties of endophytes both phenotypically and at the genome level, and keeping in mind their virulence potential, we tested whether or not these beneficial microbes can be naturally delivered by a grapevine associated insect. Our goal was to analyze if the whole endophytic community can be transferred (since we believe an inoculum ressembling the entire endophytic community could provide more benefit to the plant) and the effect that host exert on community structure so we can eventually think of using endophytes as a natural vaccination system for plants.

Our data show that the whole microbial communities are transported between plants by insect vectors. By feeding and touching the plant, insects acquire a set of microorganisms that radically differs from those they have at hatching. Insects are able to carry and transfer this set to other plants they dwell in and feed upon.

4.3.1. Structure of the community in the tested grapevine holobiont

The relative abundances of bacterial phyla varied between hosts. The control samples had a distinctively different species composition than the test samples (Figure 9 and Appendix Table 10). In the SRC, the bacterial community was mainly composed of Proteobacteria, where the most abundant classes were Beta-, Gamma- and Alpha-proteobacteria. The rest of the community was composed of Actinobacteria (a majority in the class Actinobacteria), Firmicutes (with most members affiliated with the class Bacilli), Bacteroidetes (the classes Sphingobacteriia and Spirospirae made the majority of the phylum) and, in a smaller proportion, Acidobacteria (represented only by the class Solibacteres) and Chlamydiae (represented only by the class Chlamydia). Only a small fraction of the OTUs could not be assigned to any particular taxon. The bacterial community of IN was composed mostly of Proteobacteria followed by Firmicutes, Actinobacteria, Bacteroidetes and Acidobacteria, the latter representing the least abundant phylum. The INSURF community was similar to the inner bacterial microbiota of IN with the majority of OTUs assigned to Proteobacteria, followed by Firmicutes, Actinobacteria, Bacteroidetes and Acidobacteria. In our analysis, only one OTU was exclusively associated with the INSURF samples. Using the Basic Local Alignment Search Tool (BLAST), this sequence was assigned to the Sinobacteraceae, a family that includes the closely related water-spring associated bacterium *Nevskia* sp. This sequence was never detected in any plant sample or inside the insects.

Microbiota of the SNK was mostly composed of Proteobacteria, with Beta- and Gammaproteobacteria being the most abundant classes. Deltaproteobacteria were also present and the rest of the phyla had only few representatives (Appendix Table 10). A further analysis of the endophytic community composition in the above- and below-ground compartments revealed differences in the two plant compartments. Proteobacterial OTUs in STEMSNK were highly represented (in order of abundance: Betaproteobacteria,

Gammaproteobacteria, Alphaproteobacteria and Deltaproteobacteria). Actinobacteria, Firmicutes, Acidobacteria, Bacteroidetes and Chlamydiae were less abundant.

In ROOTSNK, Proteobacteria were also the most abundant OTUs (in order of abundance: Beta-, Gamma-, Alpha- and Delta-proteobacteria). Actinobacteria, Firmicutes, Acidobacteria, Bacteroidetes, Chlamydia and the candidate clade TM6 were the least abundant. In contrast, the bacterial community of control plants (CTRLROOT, CTRLSTEM, where freshly hatched insects had fed without prior contact with SRC) was dominated by Actinobacteria, with only a small proportion of Proteobacteria (Beta- and Gamma-, but no Alpha-proteobacteria). The community in CRTLIN was also dominated by Actinobacteria.

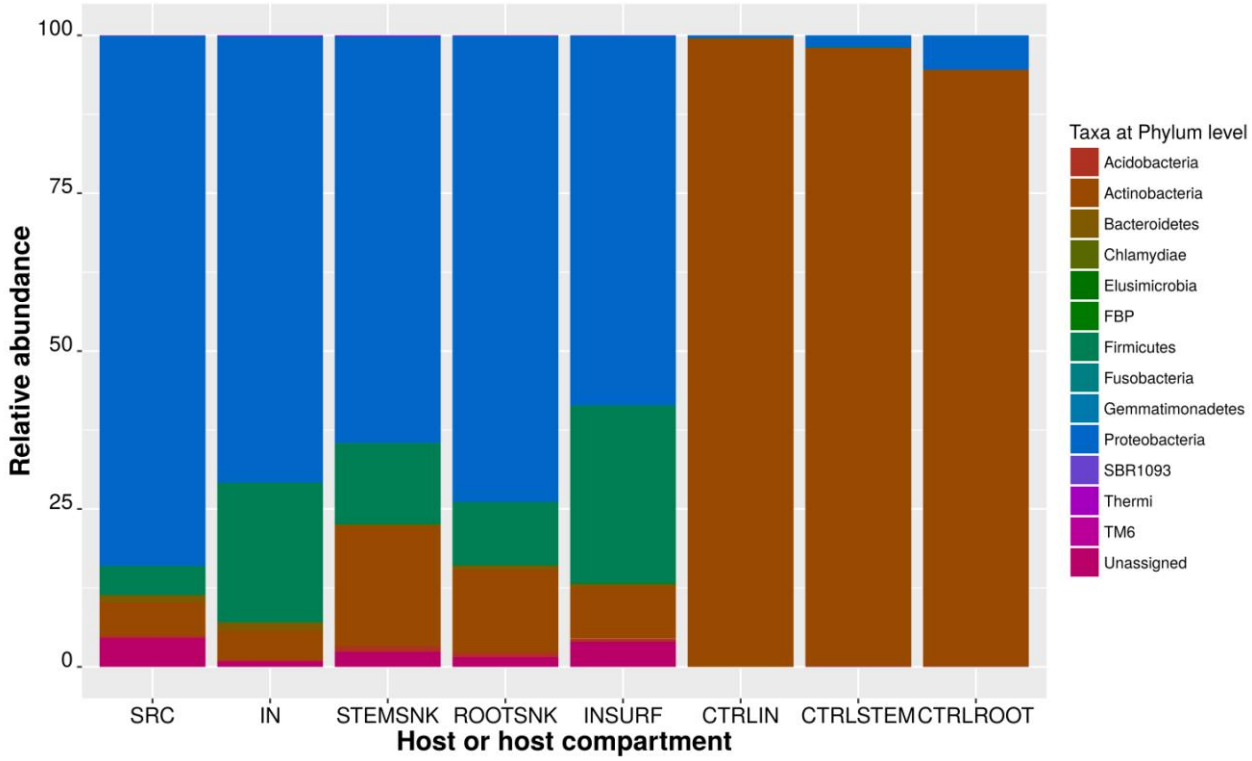


Figure 9. Relative abundance of OTUs assigned at the phylum level. Stacked bar plots represent the percentages of 216 members of the OTU Table in biom format. In the legend, unassigned correspond to OTUs whose taxonomy could not be assigned at the 97% confidence using the Greengenes database.

4.3.2. Selected endophytes are transmitted between grapevine plants

Forty OTUs were transferred by IN from SRC to SNK and found both in ROOTSNK and in STEMSNK (Table 5). These were never found in the CTRLROOT or CTRLSTEM, suggesting that they were efficiently transmitted from SRC to SNK by IN. In particular, among the sequences of the main phyla (i.e. Proteobacteria, Actinobacteria, Bacteroidetes, Chlamydiae and Firmicutes), Proteobacteria were the most highly represented taxon, where the most abundant genera were *Agrobacterium*, *Paracoccus*, *Sphingomonas*, *Erwinia*, *Pseudomonas*, *Lysobacter* and *Stenotrophomonas*.

In contrast, the most abundant phylum in CTRLIN was Actinobacteria, especially the genera *Mycobacterium*, *Gordonia*, *Nocardia*, *Rhodococcus* and *Williamsia*. CTRLROOT and CTRLSTEM hosted a community that resembled that of the CTRLIN. For example, the genus *Nocardia* was detected in all sample types, but its

prevalence was lowest in CTRLROOT samples. Likewise, *Rhodococcus* and *Aeromicrobium* were less abundant in CTRLROOT and CTRLSTEM than in CTRLIN. An exception was the genus *Williamsia*, which was more abundant in CTRLROOT and CTRLSTEM than in CTRLIN.

4.3.3. Insects change the community structure during passage from source to sink plants

To identify shifts in bacterial community composition in the transferring process, we analyzed diversity for every host type and compared their significance at a large scale (higher taxonomic hierarchy or phylum level) and in some cases at a small scale (genus level).

The largest diversity was present in ROOTSNK followed by STEMSNK. The third most diverse microbiota was that of SRC followed by IN (Figure 10). In terms of richness, a large number of new species was detected in SRC (observed species and Chao 1 index), though the number of species in IN acquired during feeding was less than a half of what it had been in the SRC (Figure 10 and Appendix Figure 6). In addition, variance within samples was larger in IN as compared to SRC, suggesting differences in species composition in each sample. In STEMSNK the richness increased considerably more than in the below-ground part of the plant, with a high variance within samples, suggesting that the community that waspreviously sTable in SRC was disturbed in the SNK after acquisition and transmission by IN.

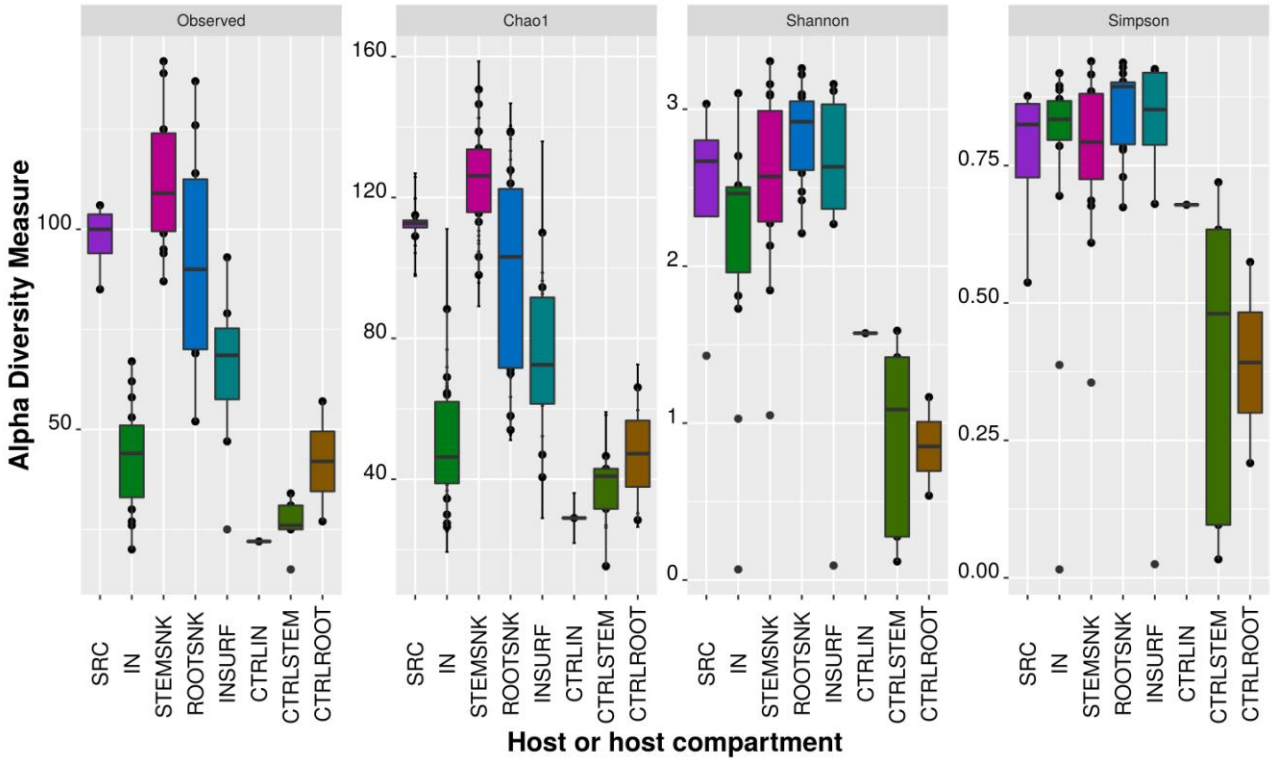


Figure 10. Alpha diversity metrics on transmission experiments. Alpha diversity indexes (Richness = Observed species and *Chao1*; Diversity = Shannon-Wiener and Simpson) were calculated on the OTU Table at the phylum level. Reads were rarefied to 1300 sequences to have an even representation of OTUs in each sample category (or host).

Table 5. OTUs transmitted from source (SRC) to sink plants (STEMSNK and ROOTSNK) by *S. titanus*

OTU	PHYLUM	CLASS	ORDER	FAMILY	GENUS	SPECIES
1	Actinobacteria	Actinobacteria	Actinomycetales	Unclassified	Unclassified	Unclassified
2	Actinobacteria	Actinobacteria	Actinomycetales	Corynebacteriaceae	Corynebacterium	durum
3	Actinobacteria	Actinobacteria	Actinomycetales	Geodermatophilaceae	Unclassified	Unclassified
4	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Kocuria	palustris
5	Actinobacteria	Actinobacteria	Actinomycetales	Micrococcaceae	Micrococcus	luteus
6	Actinobacteria	Actinobacteria	Actinomycetales	Nocardioidaceae	Unclassified	Unclassified
7	Actinobacteria	Actinobacteria	Actinomycetales	Propionibacteriaceae	Propionibacterium	Unclassified
8	Bacteroidetes	Flavobacteriia	Flavobacteriales	Flavobacteriaceae	Flavobacterium	Unclassified
9	Bacteroidetes	Saprospirae	Saprospirales	Chitinophagaceae	Sediminibacterium	Unclassified
10	Chlamydiae	Chlamydiia	Chlamydiales	Parachlamydiaceae	Unclassified	Unclassified
11	Firmicutes	Bacilli	Bacillales	Bacillaceae	Bacillus	flexus
12	Firmicutes	Bacilli	Bacillales	Staphylococcaceae	Staphylococcus	aureus
13	Firmicutes	Clostridia	Clostridiales	Tissierellaceae	Anaerococcus	Unclassified
14	Proteobacteria	Alphaproteobacteria	Rhizobiales	Phyllobacteriaceae	Unclassified	Unclassified

15	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Unclassified	Unclassified
16	Proteobacteria	Alphaproteobacteria	Rhizobiales	Rhizobiaceae	Agrobacterium	Unclassified
17	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Paracoccus	Unclassified
18	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Rhodospirillaceae	Unclassified	Unclassified
19	Proteobacteria	Alphaproteobacteria	Rickettsiales	Unclassified	Unclassified	Unclassified
20	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Unclassified	Unclassified
21	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Kaistobacter	Unclassified
22	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingomonas	Unclassified
23	Proteobacteria	Betaproteobacteria	Burkholderiales	Alcaligenaceae	Achromobacter	Unclassified
24	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	Unclassified	Unclassified
25	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	Unclassified	Unclassified
26	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	Kingella	Unclassified
27	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	Neisseria	Unclassified
28	Proteobacteria	Betaproteobacteria	Neisseriales	Neisseriaceae	Neisseria	cinerea
29	Proteobacteria	Deltaproteobacteria	Unclassified	Unclassified	Unclassified	Unclassified

30	Proteobacteria	Deltaproteobacteria	Myxococcales	0319-6G20	Unclassified	Unclassified
31	Proteobacteria	Gammaproteobacteria	Alteromonadales	Alteromonadaceae	Marinobacter	Unclassified
32	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Unclassified	Unclassified
33	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	Erwinia	Unclassified
34	Proteobacteria	Gammaproteobacteria	Legionellales	Unclassified	Unclassified	Unclassified
35	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae	Pseudomonas	nitroreducens
36	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Sinobacteraceae	Unclassified	Unclassified
37	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Luteimonas	Unclassified
38	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Lysobacter	Unclassified
39	Proteobacteria	Gammaproteobacteria	Xanthomonadales	Xanthomonadaceae	Stenotrophomonas	Unclassified
40	TM6	SJA-4	Unclassified	Unclassified	Unclassified	Unclassified

When looking at Shannon-Wiener and Simpson's indexes, diversity was found to be higher in SRC (abundance of new species is larger and even) and lower in IN (Appendix Table 11). When the community was transferred to STEMSNK, its diversity increased. The standard variation within samples increased from SRC (0.878) to IN (10.52), suggesting less evenness per sample. When passing to the SNK, the standard variation decreased (STEMSNK = 10.52 and ROOTSNK = 0.430), pointing to a recovery of the community when moving from the insect to the plant host. Although INSURF had few new species per sample, the diversity was near to that of the other samples, hinting at the possibility of an input of insect surface-associated bacteria to the endophytic community (Appendix Table 11A). The diversity of the controls differs from that of the treated samples. The lowest diversity and richness were recorded for CTRLIN, CTRLROOT and CTRLSTEM. Rarefaction curve analysis (Appendix Figure 6) confirms that the richest samples are SRC and the compartments of the sink plants (STEMSNK and ROOTSNK), followed by the INSURF microbiome and by the inner microbiome IN. CTRLROOT were the poorest samples in terms of new species, no new OTUs being discovered after a sampling effort of 1505 sequences/sampling. The rarefaction analysis shows that a large proportion of the endophytic bacterial community is represented in our analysis, since all curves showed a plateau starting at and saturated after the threshold of 1505 sequences/sampling.

Statistical analysis using non parametric t-tests on richness and diversity indexes unveiled interesting patterns that delineated the transmission pathway. With the Chao1 richness estimator, we found significant differences between SRC and IN, STEMSNK and CTRLSTEM, IN and ROOTSNK, STEMSNK and IN, INSURF and CRTLSTEM, and CTRLSTEM and ROOTSNK (Appendix Table 11B)

Statistical comparisons of the Shannon-Wiener indexes per sample group showed significant differences between CTRLSTEM and STEMSNK as well as differences between the CTRLSTEM and the ROOTSNK. This further confirmed the host effect on the bacterial community. As we expected, IN and ROOTSNK showed a significant diversity shift, suggesting that during the passage from insect to sink plants the community is shaped in a host-dependent manner. A tendency to differential diversities was detected between CTRLSTEM and ROOTSNK, although this was a borderline difference was (Appendix Table 11B).

The non-parametric t-test analysis showed that the community of INSURF and IN was mainly composed of the same taxa, with additional presence of particular groups, namely Fusobacteria and Gemmatimonadetes, on the surface of the insect. Although both sample types have overall the same taxa composition, their abundances also varied. The taxa that differed in their abundances were *Actinomyces*, *Burkholderia*, *Haemophilus parainfluenzae*, *Kocuria palustris*, *Streptococcus* sp., Neisseraceae, *Kingella* sp. and a member of the Micrococcaceae, whose taxonomy remains to be determined.

4.3.4. Endophytic community composition shifts in a host- specific manner

Differentially abundant OTUs were analyzed for each host (Figure 11). Overall, 25 taxa were differentially abundant. A strong influence of the plant host was observed, since abundance of OTUs assigned to *H. parainfluenzae*, *Kingella* sp., *Burkholderia* sp., *Kocuria palustris*, *Acinetobacter* sp., *Neisseria* sp., *Micrococcus luteus* and *Methyobacterium adhesivum* plummeted when passing from SRC to IN. We also observed how the sample type affected the relative abundances of specific OTUs.

In SRC, IN, STEMSNK, ROOTSNK and INSURF, the most abundant OTU was assigned to *Ralstonia* sp., with the highest number of sequences in the STEMSNK. In CTRLIN, CTRLROOT and CTRLSTEM, the most

abundant OTU was assigned to *Mycobacterium* sp., with the highest number of sequences in the CTRLSTEM (Appendix Table 12).

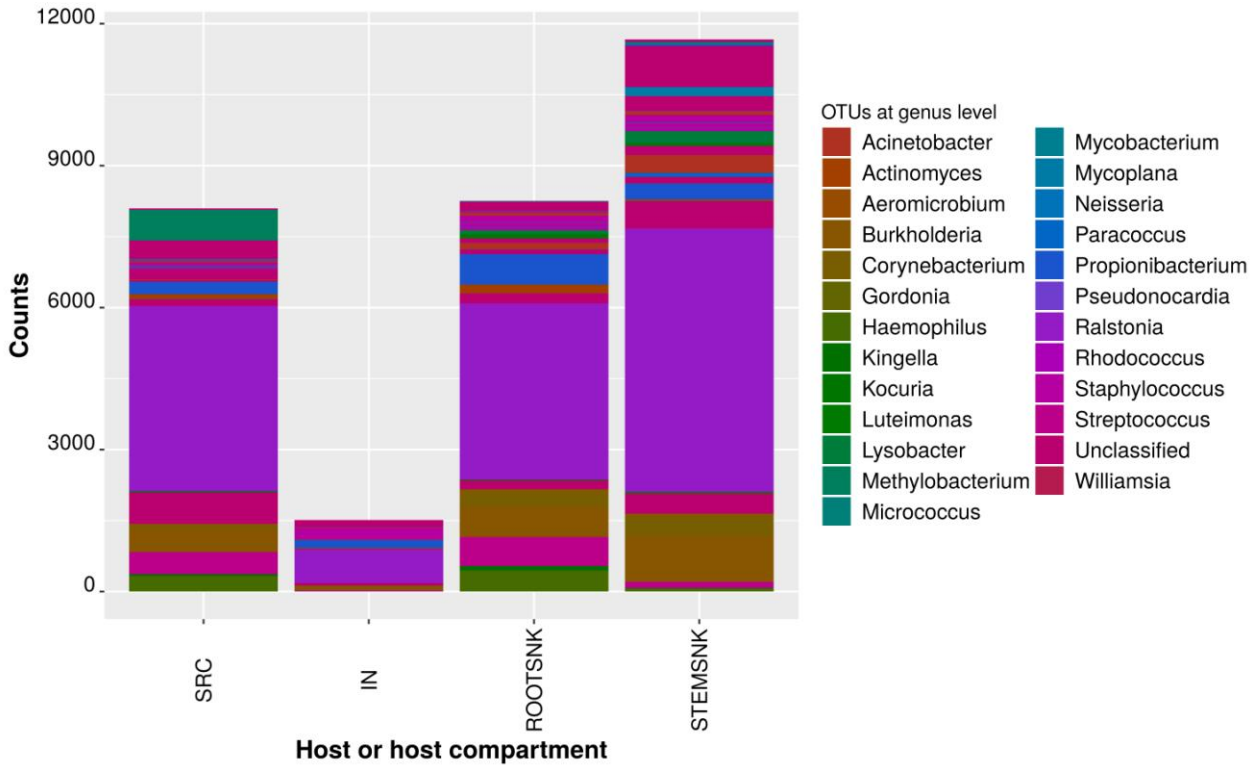


Figure 11. Comparisons of statistically significant abundances per sample category in a rarified OTU Table. Frequencies per sample categories with non-normal distribution were compared using the Kruskal-Wallis test as implemented in QIIME. Only SRC, IN, STEMSNK and ROOTSNK were analyzed. The probability of finding at least one mean significantly differing from the others was corrected using the False Discovery Rate (FDR) and the Bonferroni procedures, as implemented in QIIME.

In contrast, in SRC, IN, STEMSNK and INSURF, the least abundant OTU was *Paracoccus* sp. In ROOTSNK the least abundant OTU was assigned to *Staphylococcus* sp. In CTRLIN, CTRLROOT and CTRLSTEM, the least abundant OTU was also *Paracoccus* sp. Comparison between samples indicated that bacterial communities are shaped by the insects or the plant and adapt to the hosts, altering the relative frequencies of key taxa). The analysis of principal coordinates showed that samples from SRC and ROOTSNK/STEMSNK grouped together, while control samples (CTRLROOT and CTRLSTEM as well as CTRLIN) formed a separate group (Figure 12). The IN samples separated from the rest of the groups. The microbiota of SRC and SNK showed that Proteobacteria, Actinobacteria and Bacteroidetes group together with few Chlamydia sequences and some members of the TM6 clade of amoebal symbionts. In the CTRLIN, CTRLROOT and CTRLSTEM, the Actinobacteria form a separate group in the plot, suggesting differentiation from the rest of the community. The community in IN, predominantly comprising Proteobacteria, groups together with Fusobacteria. The surface microbiota of IN appear separate from IN and the plant samples (ROOTNSK and SRC) and are dominated by Proteobacteria.

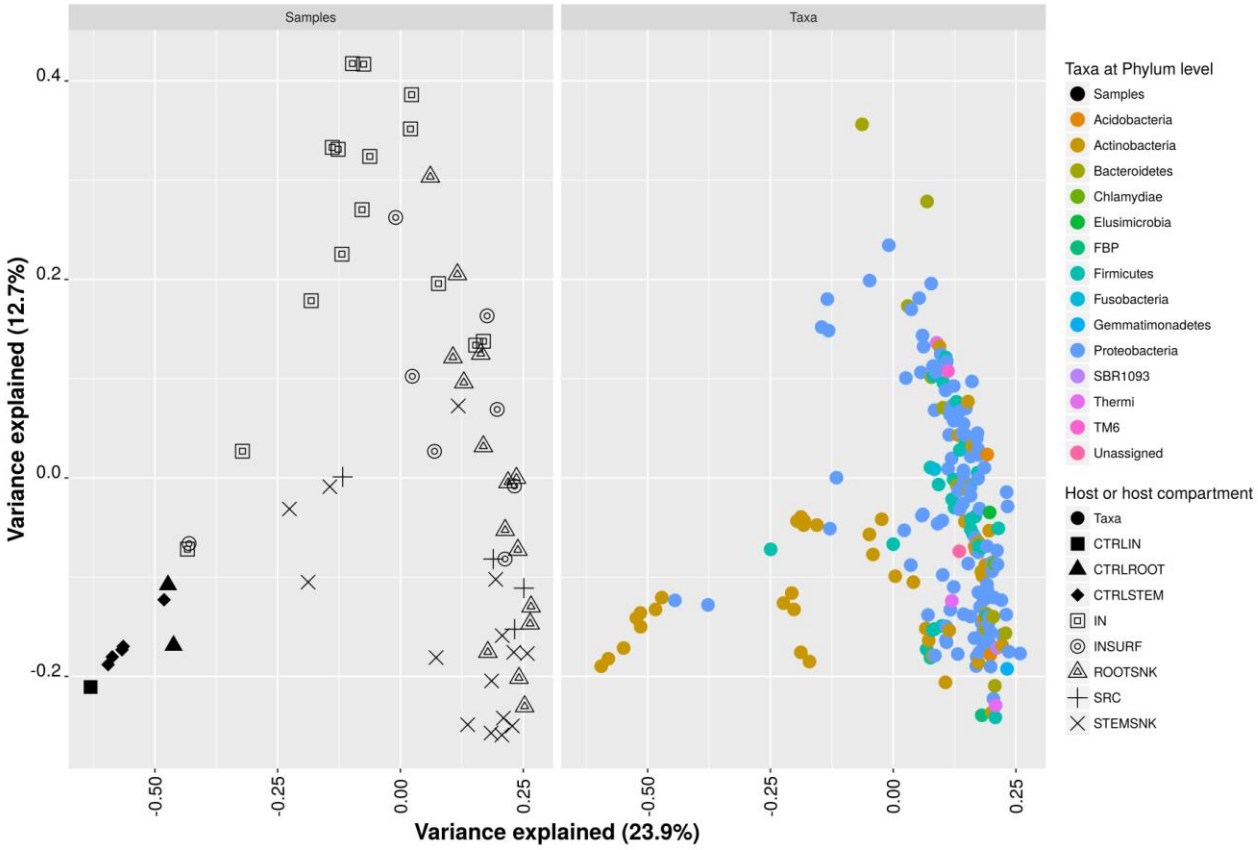


Figure 12. Principal coordinates analysis of a rarefied OTU Table at phylum level. Dissimilarities were calculated using the Bray-Curtis distance estimator. Variance per axis is presented as percentage

We further analyzed the dynamics of bacterial abundance fluctuation using ternary composition plots with bacterial endophytic communities of SRC, IN, STEMSNK and ROOTSNK. This analysis suggests that some taxa are overgrown by others, in a host-dependent manner (Appendix Figures 7-8). In the diagrams, intersection of the perpendicular segments shows several phyla simultaneously. For example, at the tip of the triangle, a high density of OTUs is depicted (Appendix Figure 7). These OTUs are represented in lower abundances in SRC and IN (between 1 and 20%), while the same OTUs are in high abundance (80-100%) in the STEMSNK. Since diameter of the circles represents the relative abundance of OTUs in the three samples analyzed, the plot suggests that Actinobacteria are the taxon most commonly found in all samples (biggest circle is at coordinates 20% SRC, 1% IN and 80% STEMSNK). In addition, a gradient can be seen where the community dominated by Proteobacteria and Actinobacteria in SRC and IN gradually changes to be composed of other groups like Firmicutes, Fusobacteria and Acidobacteria. This gradient is easier to spot at the top vertex of the triangle, where a high density of different OTUs is seen at 100% abundance in the STEMSNK, as compared for example to the lower left vertex, where the community was mostly represented by Actinobacteria and Proteobacteria. In this gradient, emergence of Fusobacteria, Firmicutes and Bacteroidetes (although not highly represented in each sample) delineates a possible displacement of the initial-community dominated by Proteobacteria in SRC to a more diverse community in STEMSNK. Similar

results were obtained for ROOTSNK, although Fusobacteria and Firmicutes had already been found in the IN sample (Appendix Figure 7, lower right vertex of the triangle).

4.3.5. Endophytes are acquired by insects through feeding and delivered to the stems of grapevine plants

The qPCR analysis detected eGFP-tagged bacteria in both STEMSNK and ROOTSNK as well as in the IN vectors, confirming direct transmission of endophytes from insect to plant through feeding (Figure 13). In samples inoculated with *E. ludwigii* EnVs6 (pMP4655), we consistently detected more than 104 eGFP gene copies/g of plant tissue in ROOTSNK, STEMSNK and IN. eGFP quantification was lower (between 10^4 and 10^6 gene copies/g of plant tissue) in samples inoculated with *E. ludwigii* EnVs2 in ROOTSNK. However, the bacterium reached titers between 10^7 and 10^9 eGFP copies/insect in the STEMSNK samples and up to 10^7 in IN samples. In contrast, endophytic *Pantoea vagans* PaVv9 was only partially transmitted. Quantification of gene copy numbers in SNK showed lower colonization levels of this endophyte, reaching only 10^2 eGFP gene copies per gram of plant tissue in one STEMSNK, and being undetecTable in one replicate of the ROOTSNK. The standard of 10^5 gene copies per gram of plant tissue suggests complete amplification, confirming that the low numbers were the result of a poor transmission by the insect and not an artifact of the PCR. We also performed endpoint PCR amplification in all the samples with specific primers for 16SrDNA genes of endosymbionts of *S. titanus*, i.e. *Cardinium* sp. and *Asaia* sp. We did detect *Asaia* sp. in the IN samples, though *Cardinium* sp. was not detected in either SRC, IN or SNK.

4.3.6. Remarks

Our experimental setup allowed us to confirm that *S. titanus* can transmit a wide variety of bacterial endophytes.

When bacterial sequences were detected in only one of the hosts, *i.e.* SRC, IN and SNK, we hypothesize that they were part of host microbiota, which was not transferred. On the contrary, when sequences were present in SRC, IN and SNK, we hypothesize that the bacterial communities had been efficiently transferred from SRC to SNK through IN, in a contact-dependent manner. We also controlled for species transmitted from IN's own microbiomes by using the CTRLIN that had no contact with SRC and by sequencing the INSURF, detecting the community adhered to the insect's surface. When sequences were detected in the controls (CTRLSNK and CTRLIN) as well as in SRC, SNK and IN, we speculate that these species belong to the insect's microbiota and can be found in any of the samples that had contact with the insect.

We detected an overall fluctuation of diversity and species richness across samples. In Figure 10, we show that richness drops following transfer from plant to vector, plausibly as an effect exerted by the host. It is well known that microbial communities are adapted to very particular micro-environments in the host and this is one of the factors that shape microbiota structure. This effect can be seen when comparing microbiotas of non-phylogenetically related hosts, but also in closely related hosts, where it has a particular community structure.

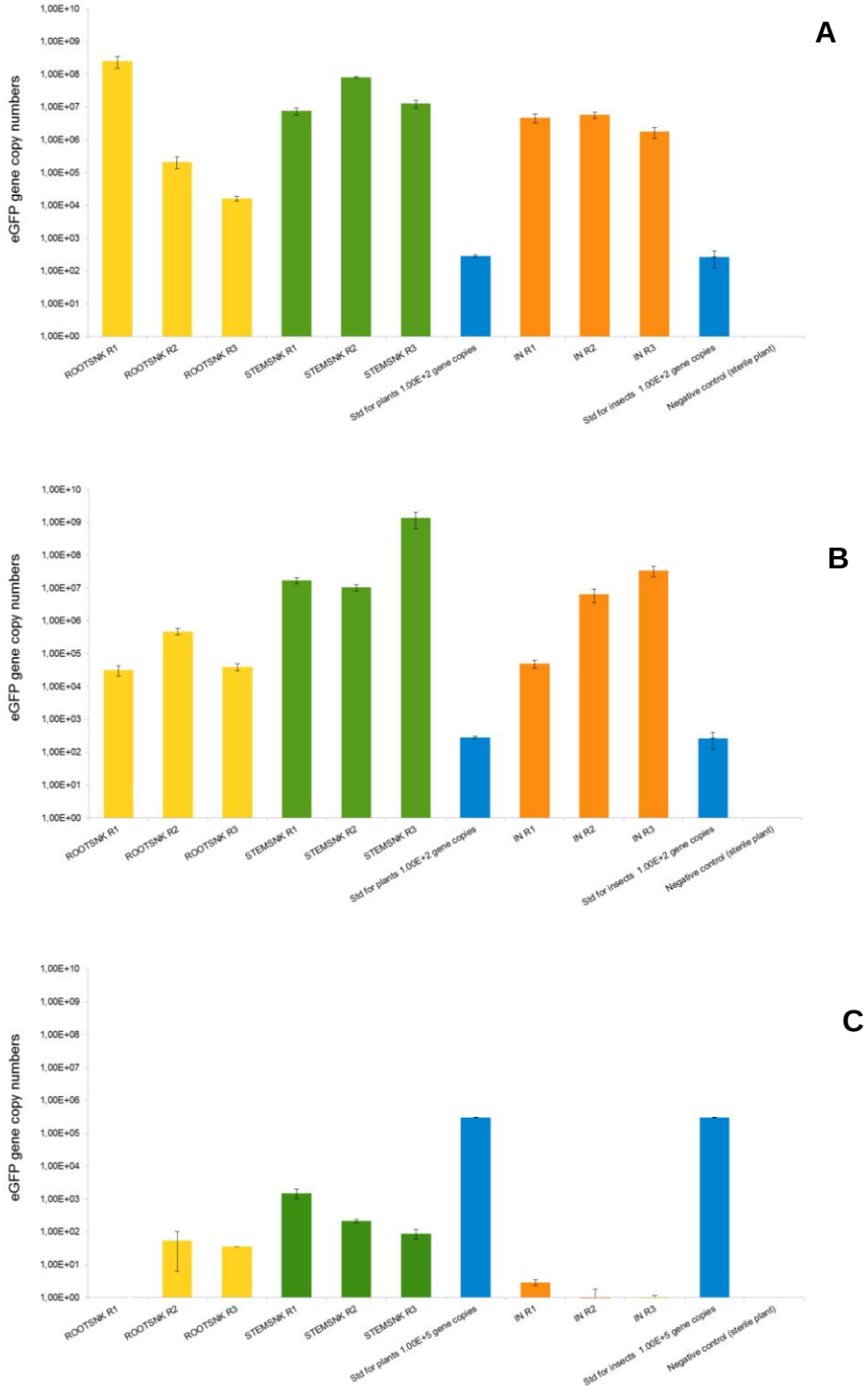


Figure 13. Quantification of eGFP gene copy numbers in the plants and insects used in transmission experiments. Gene copy numbers in the eGFP-bearing plasmid pMP4655 were quantified using qPCR with the second derivate method. For plants, the values obtained were benchmarked against the fresh weight (g) of samples; for insects, the copy numbers per individual are shown. Bars represent the mean of five biological replicates per treatment. Error bars represent the standard deviation of the mean. **A.** quantification of eGFP in *Enterobacter ludwigii* EnVs2. **B**. quantification of eGFP in *Enterobacter ludwigii* EnVs6. **C**. quantification of eGFP in *Pantoea vagans* PaVv9.

Studies have shown that different pig breeds harbor differential microbiota structures and that these can be transferred to non-related animals, *e.g.* mice, which in turn will also modify community structure, highlighting the specificity across hosts (Diao et al., 2016). Greater variance in richness estimators (*Chao1* and observed species) denotes variable content of species per sample (Fig. 10). In diversity estimators (Shannon-Wiener and Simpson), greater variance represents the non-dominance (equal distribution) of species, while small variances per sample represent dominance (unequal distribution) of a few species in the community.

These effects can be easily seen in terms of Simpson's diversity estimator, where SRC's microbiota had a higher variance, which decreased as it was acquired by IN. In STEMSNK and ROOTSNK, variance of the estimator increased, once again suggesting the non-dominance of particular species. When analyzing these fluctuations as a composite, we are observing an "expansion and contraction" phenomenon of the bacterial endophytic community where changing hosts alters the community structure and possibly changes in functional groups. From this perspective, species dominance might hint at a bottleneck for specialists (bacteria with restricted enzymatic capabilities) that might be increased and dominate when inhabiting the IN. When inhabiting the plant hosts (SRC and SNK), generalists (*e.g.* bacteria with a wider range of enzymes capable of using resources in the ecosystem) are evenly distributed.

Besides a plausible role of nutritional factors in community shaping, when passing from plant to the insect host, we presume that niche occupation might be a trigger for competition among bacteria and thus affect diversity. In our experiments, niches in SRC's microbiota may have overlapped with those of IN's microbiota resulting in competitive exclusion. This could result in a decrease in the number of species from SRC to IN when abundance of some genera drastically changed from plant to insect (Figure 11 and Appendix Table1). Among those microorganisms whose abundance decreased, we found plant- and human-associated bacteria that have been previously identified in the microbiome of grapevine (Yousaf et al., 2014). We believe that these bacteria, might behave as drivers of diversity with roles in directing community structure during host colonization, given the specificity of their change in abundance and because they harbor genetic determinants that make them good competitors.

As opposed to the SRC, SNK and IN, CTRLSTEM, CTRLROOT and CTRLIN clearly differed in terms of community composition, where Actinobacteria and Proteobacteria were completely dominant (Figure 1 and Figure 3). Just a low number of the transferred endophytes found in SRC, SNK and IN were found in the controls.

We have also shown a distinction between the insect's microbiota and the plant's endophytic community and propose differentiation according to the host (Figure 12). Shifts in the bacterial community are associated with a gradient where some of the taxonomical groups have more weight during transmission than the others (Appendix Figures 7 and 8). Such a differentiation suggests that some taxa are more proliferous in particular hosts. This is of extreme importance in terms of possible endophytic microbiota manipulation by engineering. If some of the endophytes (for example those assigned to Bacteroidetes in IN or Acidobacteria in SRC) can grow better in certain types of hosts, we might drive our efforts to enrich and improve these groups in those hosts and experiment to further analyze their plant protection properties for use in agriculture. .

We also highlight the selectivity of the insect as a vector of endophytes. Our qPCR experiments (Figure 13) suggest that endophytes like *E. ludwigii* EnVs6 and *Enterobacter* sp. EnVs2 are effectively vectored by the insect from an exogenous source. However for the endophyte *P. vagans* PaVv9, the transmission was

inefficient. It is possible that only few members of the endophytic community that are acquired by the insect will survive in this host. This kind of selectivity has been observed in pathogens of grapevine. For example, *S. titanus* is able to transfer the FDP more efficiently than any other leafhopper vector (Chuche and Thiéry, 2014). This, together with the high grapevine specificity could make *S. titanus* an ideal candidate for endophytic delivery in future therapeutic applications in the vineyards.

4.4. Grapevine colonization by endophytic bacteria shifts secondary metabolism and suggests activation of defense pathways

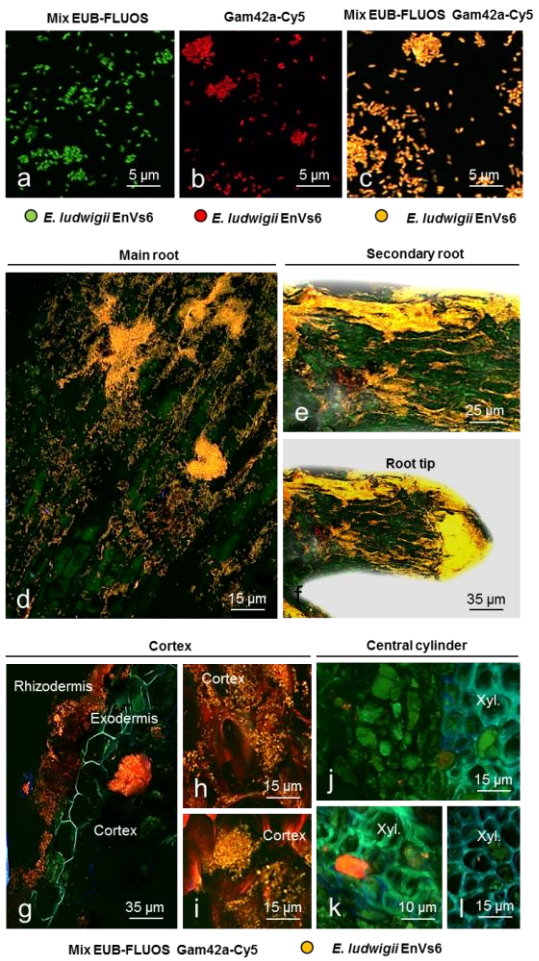
Delivery of endophytes through insects is an efficient process, as shown above. What could eventually happen when endophytes reach the plant (i.e. one they are delivered)? Understanding of the colonization process and its effect on plants homeostasis could allow the conception of ways to enginerize transmissible endophytic communities so that endophytes that are transported by the insect could be also effective colonizers, not affecting plant quality or health.

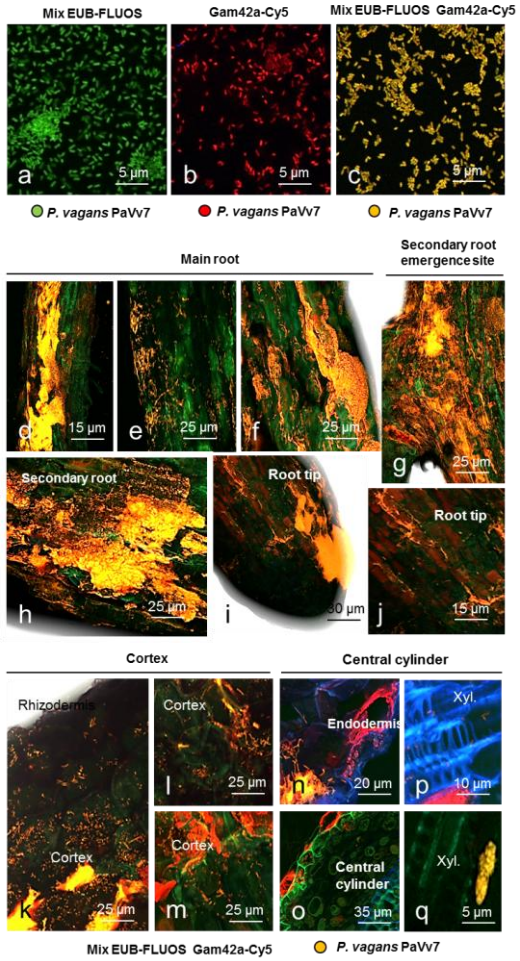
4.4.1. Colonization of bacteria visualized through DOPE-FISH

DOPE-FISH microscopy allowed visualizing *E. ludwigii* EnVs6 using EUBmix, Gam42a or probe combination (Fig. 14a-c). In plantlets, strain EnVs6 was observed on the rhizoplane of the main root (Fig. 14d), in secondary roots (Fig. 14e), as well as at the root tip level (Fig. 14f). Strain EnVs6 was detected as an endophyte in the cortex (Fig. 14g-i), as well as inside the central cylinder up to the xylem vessels (Fig. 2j-l) while no bacterial colonization was recorded in the aerial plant parts (data not shown). Similar experiments resulted in detection of *P. vagans* PaVv7 in pure cultures (Fig. 15a-c) as well as on the rhizoplane of grapevine plantlets (Fig. 15d-j) and inside plant tissues (Fig. 15k-q). Strain PaVv7 was detected particularly as colonizing the rhizoplane of plantlets at the main root level (Fig. 15d-f), secondary root (Fig. 15g-h) as well as at the root tip level (Fig. 15i-j) and bacteria were visualized as single cells (Fig. 15e-j) or colonizing the whole outline of some rhizodermal cells (Fig. 15d, and 15f-i). Bacterial aggregates as well as microcolonies were recorded in some plant parts (see Fig. 15f and 15h-i). Endophytism by strain PaVv7 was observed in the cortex (Fig. 15k-m), as well as inside the central cylinder (Fig. 15n-q) and up to the xylem vessels (Fig. 15p-q). Similarly to strain EnVs6, no endophytic colonization in the aerial plant parts was detected on all the examined plantlets (data not shown).

Strain *S. phyllosphaerae* SpVs6 was also hybridized and examined as pure culture or during the colonization of grapevine plantlets except that only EUBmix probes were used. Fig. 16a shows the pure culture of strain SpVs6. Interestingly, strain SpVs6 was detected as colonizing the rhizoplane of the plantlets at the main root level (Fig. 16b-c), secondary root (Fig. 16d-e) or slightly on root tips (Fig. 16f) but was never detected as an endophyte inside roots as well as inside tissues of shoots.

As expected, in non-inoculated control plants, no bacteria were visualized on the root surface (Appendix Figure 9a-c) as well as inside root internal tissues (Appendix Figure 9d). Additional controls with NONEUB probe did not result in visualization of bacteria on samples of grapevine inoculated with EnVs6 (Appendix Figure 6a-b), PaVv7 (Appendix Figure 6c-d), SpVs6 (Appendix Figure 6e-f) or control treatment (Appendix Figure 6g-h).





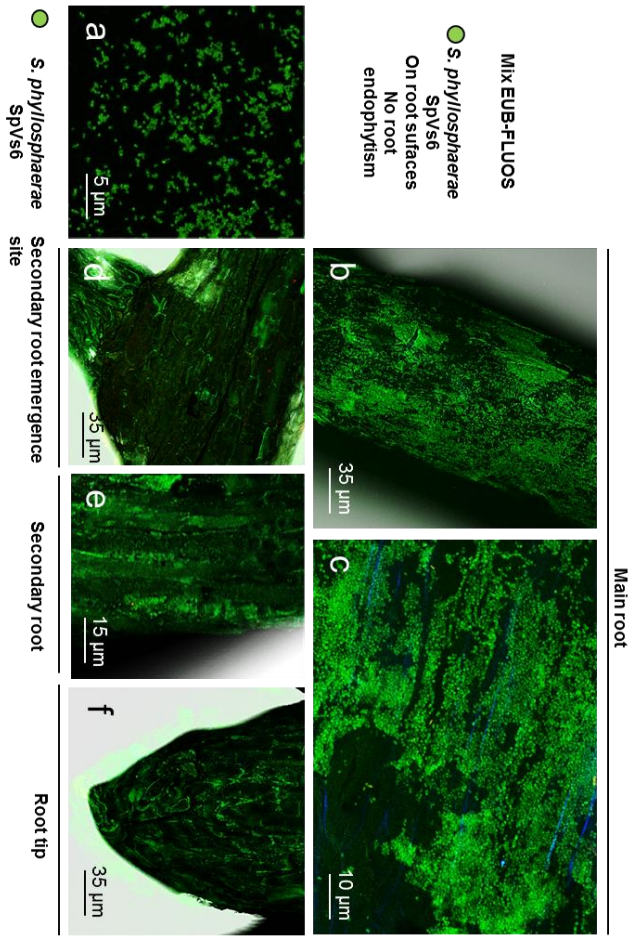


Figure 14 (left). Visualization of *Enterobacter ludwigii* EnVs6 in cell cultures (a-c) and in *Vitis vinifera* L. (d-l) by DOPE-FISH. Cultured bacterial cells of strain EnVs6 tagged with a) the EUBmix probes b) the Gam42a probe c) a cocktail of probes EUBmix and Gam42a. Bacterial cells of strain EnVs6 hybridized with the EUBmix and Gam42a probes colonizing *Vitis vinifera* L. on d) main root e) secondary root f) secondary root tip g) rhizodermis, exodermis and cortex h) cortex with disseminated bacterial cells i) cortex with aggregated cells j) xylem with few bacterial cells k) aggregated cells in xylem vessels l) bacterial cells on xylem vessels**.**

Figure 15 (middle). Visualization of *Pantoea vagans* PaVv7 in cell cultures (a-c) and in *Vitis vinifera* L. (d-q) by DOPE-FISH. Cultured bacterial cells of strain PaVv7 tagged with a) Mix EUB probes b) Gam42a probe c) a mix of probes EUB and Gam42a . Bacterial cells of strain PaVv7 tagged with EUBmix and Gam42a probes colonizing *V. Vinifera* L. on d- f) main root g) secondary root emergence site h) secondary root showing biofilm structures i) root tip j) root tip close-up k-m) cortex with aggregated cells n) endodermis with microcolony o) cortex parenchyma and central cylinder p) xylem vessels with dispersed bacterial cells q) microcolony on xylem**.**

Figure 16 (right).Visualization of *Sphingomonas phyllosphaerae* SpVs6 in cell cultures a) and in *Vitis vinifera* L. b-f) by DOPE-FISH. Cultured bacterial cells of strain SpVs6 tagged with EUBmix probes a) Bacterial cells of strain SpVs6 tagged with Mix EUB colonizing *V. Vinifera* L. on b-c) main root, d) secondary root emergence site, e) secondary root showing dispersed bacterial cells f) root tip.

4.4.2. Metabolome analysis reveals effect of endophytic colonization on plant's metabolism

We screened for the production of 56 secondary metabolites in inoculated grapevines in two independent experiments. Of these compounds, 46.4% belonged to the flavonoid group while 26.8% were stilbenes, including several stilbenoids and related compounds. The remaining ca.27 % of compounds is organic (hydroxycinnamic) acids (Appendix Table 13). Several metabolites were present in concentrations below the detection limit before and after inoculation with strain EnVs6, including vanillin, vanillic acid and esculin (Figure 17). Contrastingly, metabolites like caftaric, fertaric and *trans*-coutaric acids reached higher concentrations.

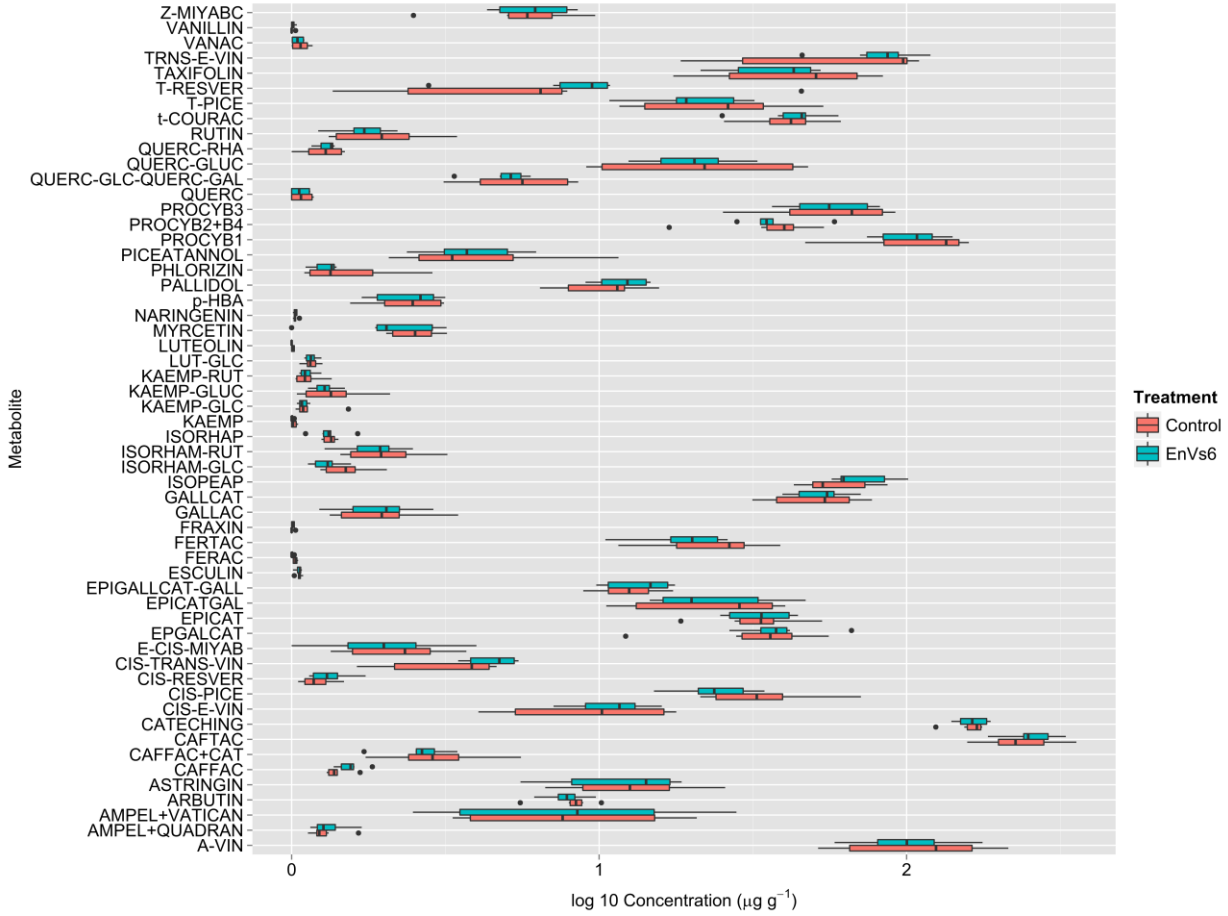


Figure 17. Concentration of metabolites in control and endophyte-inoculated grapevine plants. The boxplot represents the average log10 concentration of two independent experiments. Bars are the standard deviation (SD) of 3 replicates per treatment in two independent experiments. Alist of names corresponding to the abbreviations of metabolites in the figure, is found in Appendix Table 13.

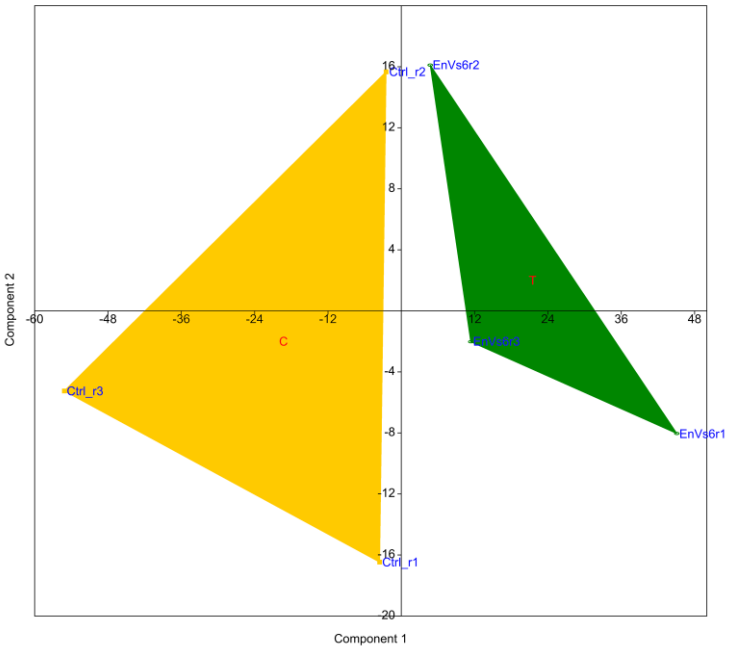


Figure 18. PCA of the first comparison between plants inoculated with strain EnVs6 and controls. Ctrl r1, r2, r3 represent grapevines inoculated with *E. coli* DH5α (green plot). EnVs6 r1, r2, r3 represent grapevines inoculated with *Enterobacter ludwigii* EnVs6 (yellow plot). C= Control; T= Treatment.

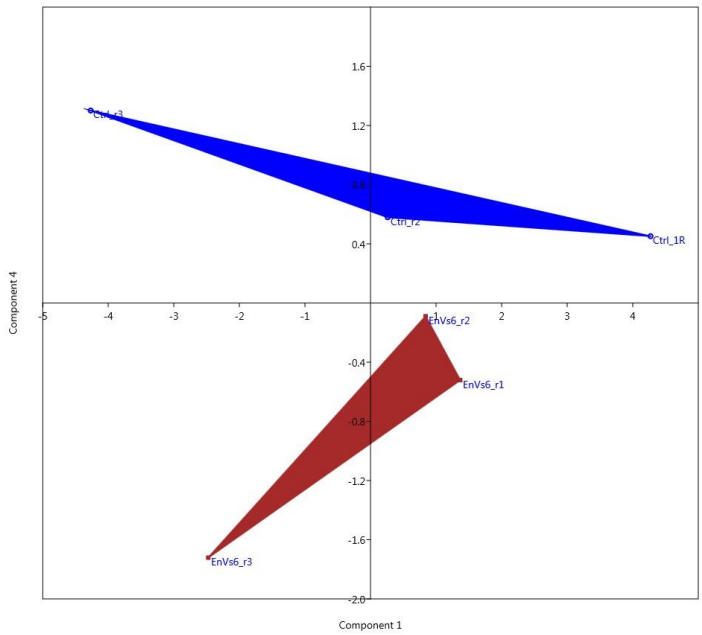


Figure 19. PCA of the second comparison between plants inoculated with strain EnVs6 and controls. Ctrl r1, r2, r3 represent grapevines inoculated with *E. coli* DH5α (blue plot). EnVs6 r1, r2, r3 represent grapevines inoculated with *Enterobacter ludwigii* EnVs6 (red plot). C= Control; T= Treatment.

Multivariate analysis indicated that data points representing the samples in the two experiments clustered independently (data not shown). A one-way analysis of similarity (ANOSIM) confirmed the difference between experiments to be statistically significant (p=0.01).

PCA on samples from each of the experiments indicated that treated samples clustered separately from controls. Figure 18 shows that plants treated with EnVs6 and control plants clustered differentially, with points representing samples from the same treatment separated along the main component that explained 79.2% of the total variance. Despite the visualized clustering using PCA, ANOSIM did not detect significant differences between the treatment and control group (R=0.074). The second experiment (Figure 19) confirmed the effects of endophytic colonization on grapevine's secondary metabolites. PCA indicated separation of the treated samples from the controls.

A two-way ANOVA for both experiments revealed a significant difference between control plants and those inoculated with strain EnVs6 ($F_{(1.560)} = 38.87$; $p < 0.0001$).

Multiple t-tests outlined a significant difference in the concentration of specific metabolites and suggested selective accumulation or depletion of metabolites following colonization. While vanillic acid accumulated in the EnVs6-inoculated plants (p= 0.00003661525), esculin (p= 0.000007356803), catechin (p= 0.000100992), kampferol (p=0.000000002935455), arbutin (p= 0.00001696334), astringin (p=0.00134368),

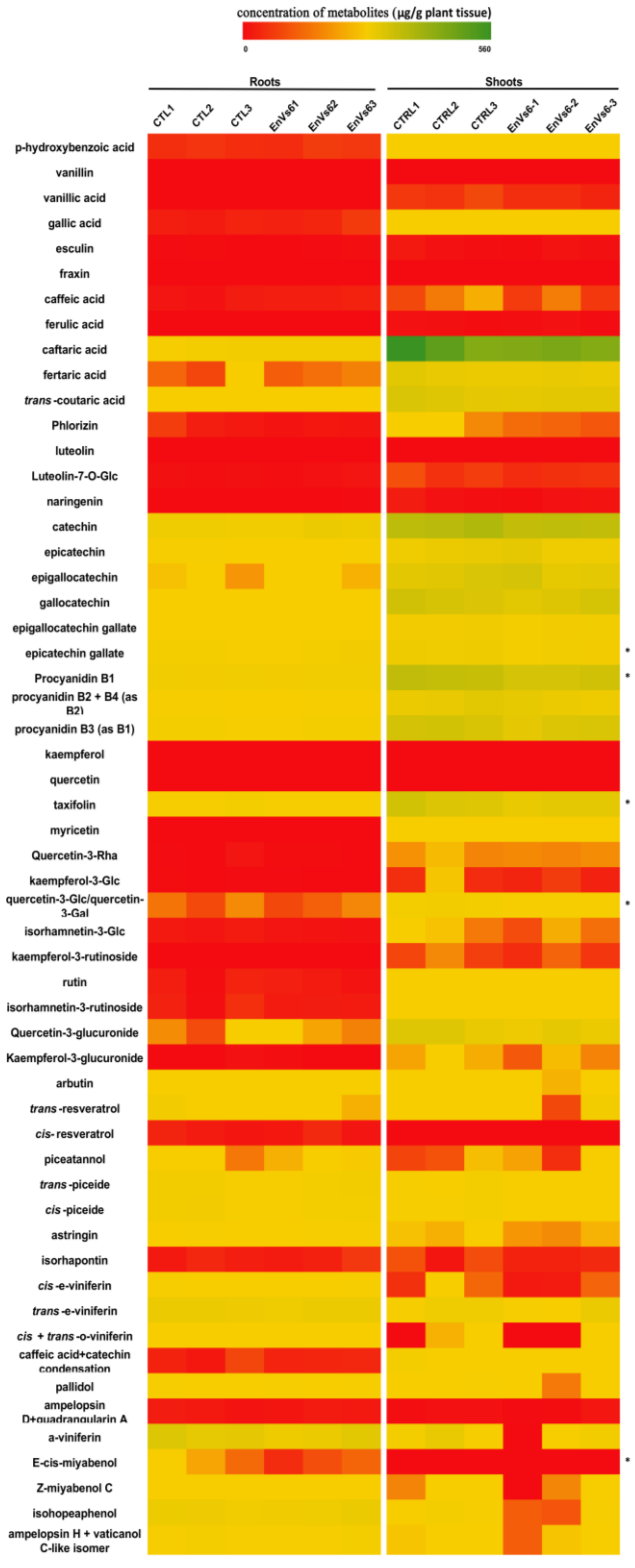


Figure 20. Heatmap showing distribution and shift of concentrations of metabolites in roots and stems of *V. Vinifera* L. Concentration of metabolites are depicted from below detection limit (red) through highest concentration (green). Asterisk denote significant differences in the concentration of control and treated organs, in a Two way ANOVA at α=0.05. CTRL: Plants inoculated with *E. coli* DH5α. R: Roots. S: Stems. Replicates are denoted by numbers 1-3.

pallidol (p=0.00130449), ampelopsin D-quadrangularin A (p=0.000439665) and isohopeaphenol (p=0.00015217) were significantly less concentrated after the treatment with this bacterial strain.

Post inoculation, several metabolites were differentially distributed in the roots and in the stems as shown in Figure 20. Roots and stems did show significant differences in the concentration of metabolites between control and treated plants. This effect was more evident in stems. Epicatechin gallate (p=0.0131), Procyanidin B1 (p=0.0029), taxifolin (p=0.0424) and the sum of quercetin-3- glucoside and quercetin-3-galactoside (p=0.0474) had a significant decrease in above ground parts after inoculation with the enterobacterium EnVs6, as compared to the control plants. A general linear model (GLM) showed the effect of plant organ and treatment in the shifts of concentration for particular metabolites. Treatment affected differently epicatechin gallate (p=0.009) accumulation in roots and stems, with major shifts in stems of plants inoculated with EnVs6; for procyanidin B1 (p<0.005) treatment affected both root and stem separately, showing its differential accumulation in both organs independently; treatment also had an effect on taxifolin (p=0.024) the sum of quercetin-3-glucoside and quercetin-3-galactoside (p= 0.024) and E-cis-miyabenol (p=0.043) accumulation in stems.

Analysis also suggested an accumulation of procyanidin B3 depending of plant organ and treatment (although this observation was not supported by a significant p= 0.051). These data suggest that both root and stem differentially accumulate the metabolite after treatment with strain EnVs6. Concentration of

caftaric acid, luteolin-7-O-glucoside and *trans*-coutaric acid suffered minor changes in concentration after inoculation with EnVs6. As expected, the concentration of metabolites in roots and stems was significantly different, with a higher concentration of metabolites in the stems as shown in Figure 20.

4.4.3. Remarks

Our results revealed the strategies for root penetration and provided a high resolution view of their tropism for root tissues. Our data also confirm the existence of a metabolic signature associated with the inoculation of grapevine with strain *Enterobacter ludwigii* EnVs6. We observed that the three strains originally isolated from the grapevine endosphere were capable of colonizing the surface of plant roots, each one with a specific pattern. Two of the endophytes used in our study (EnVs6 and PaVv7) were competent for root endophytism in the conditions tested, while one (strain SpVs6) was able to colonize only root surfaces.

Studies on endophyte colonization show that entry to the aerial parts of the plant can happen in the stomata and the hydathodes while colonization of the inner tissues of roots is usually linked to the emergence sites of secondary roots, root tip, root hairs, and by passing between the cells at other zones (Huang 1986; Sturz et al. 2000; Hardoim et al. 2008). Our observations corroborate entry of our strains through some of these points and further reinforce the notion that these sites are common colonization zones for endophytes. We believe that, as previously suggested (Burdman et al. 1999), bacterial surface-associated molecules and plant receptors play a key role in the attachment to and recognition of these colonization sites. Root endophytes also colonized parts of the xylem, which suggests that grapevine endophytes might move from the sites of attachment to internal plant tissues, and eventually spread through the plant organs as shown in Figure 14 g-l and 15 k-q. Our experiments also confirm the colonization of the xylem vessels as it has been long been known for other endophytic bacteria such as *Herbaspirillum seropedicae* (James et al., 1997). Gyaneshwar and co-workers (2001) also showed that rice is vastly colonized by the endophytic diazotroph *Serratia marcescens* strain IRBG500. This bacterium is able to colonize roots and move towards the above-ground parts by efficiently colonizing the xylem vessels. Compant et al in 2005 showed that the endophytic bacterium *Burkholderia phytofirmans* PsJN is capable of colonizing xylem vessels, cortical cells and the endodermis in primary roots of *Vitis vinifera,* similar to what we saw in our isoaltes.

We also inquired whether inoculation of plants with bacterial endophytes might lead to characteristic metabolic changes in the plant, unique for endophytic colonization. Based on the fact that some of the metabolites whose concentration shifted during the interaction endophyte-grapevine are phenylpropanoids (whose antioxidant activity has been demonstrated previously) we propose that changes observed in the metabolic profile of plants inoculated with an endophytic bacterium are an analogous response to that of plants colonizaed by fungi. The symbiosis between fungal endophytes and plants relies on the production of fungal reactive oxygen species (ROS) that mediate hyphal tip growth and interkingdom crosstalk to keep a fungal non-pathogenic state (Kogel et al. 2006; Tanaka et al. 2006). During the interaction with endophytic fungi, the content of antioxidant substances (mostly phenolic compounds) in the plant can be altered as shown previously (Malinowski et al. 1998), because plants respond to the ROS produced by the endophytes but also because the plant itself initiates an oxidative burst on the colonizing microorganism (White and

Torres 2010). Our findings suggest that the association between endophytic bacteria and grapevine is characterized by a strong metabolic cross-talk where the plant responds to bacterial colonization by shifting the concentration of specific metabolites. The fact that the increase or depletion of several metabolites, did not affect the colonization by bacterial endophytes, suggests that this shifts in the metabolic profile of the plant favor the colonization by these microorganisms. Further investigations should be done in which mutant plants in key genes for some of the pathways involved in the metabolic signature, interact with endophytes to deeply evaluate the meaning of this phenomenon in the symbiosis process.

In our experimental setting, we observed mainly an increase of vanillic acid while the compounds esculin, catechin, kampferol, arbutin, astringin, pallidol, ampelopsin and isohopeaphenol decreased in concentration after endophytic inoculation. We hypothesize that this metabolic signature is common in several examples of plant colonizing baceria. In agrobacteria-plant symbiosis, microorganisms are able to recognize signals from the host that include vanillin, guaiacol, sinapinic acid and several other phenolic compounds (Winans 1992). During colonization, agrobacteria are able to manipulate the plant cells and induce the synthesis secondary amines known as opines. In this way, growth of agrobacterial populations (and no other population of co-inoculating microorganism) will be supported by plants cells (McCullen and Binns 2006). Finally, in actinorhizal symbiosis, *Frankia* species and dicotyledonous plants are subject to an intense chemical exchange (Capoen et al. 2009; Perrine-Walker et al. 2010). The bacterium is capable of producing auxins and to sense isoflavonols to achieve colonization of the plant (Hocher et al. 2011).

Most of phenolic compounds function as phytoalexins, substances involved in protection and antibiosis against plant pathogens. It is therefore plausible that these molecules are related to the non-self recognition of colonizing organisms in different symbiotic scenarios. For example, our findings show that vanillic acid accumulates in the plant after inoculation with the endophytic strain EnVs6. Previous work has demonstrated the role of vanillic acid in the activation of root microbiota. The proposed mechanisms for this substance includes and enhancement of the production of antifungal substances in the bacterial community of the rhizosphere (Jousset et al. 2010), we presume thus that the accumulation of the metabolite during endophytic colonization might be related with an activation mechanism whereby the colonizing endophyte recognizes the molecule and initiates colonization. This last biological question should be later on addressed through experimental approaches.

On the other hand, some of the metabolites were less concentrated in plants treated with the enterobacterial strain EnVs6. Our results are in agreement with previous experiments showing that these metabolites are acting as mediators of symbiosis. In one such case, kampferol has been studied as an anti-nodulation substance in *Azospirillum* (Zhang et al. 2009). The decrease in the concentration of this metabolite during the endophytic colonization in grapevine might also be related to a down-regulation in the synthesis of such anti-symbiosis molecules during colonization, as a means of facilitating endo-symbiosis. In a similar way, arbutin and esculin (whose concentration was reduced in plants treated with strain EnVs6) may also be depleted for the same purpose, as suggested in other experimental conditions (Mo and Gross 1991). A role of the accumulation of ampelopsin D and quadrangularin A in the plant against *Plasmopara viticola* has been recently established suggesting that increase in its concentration acts as a defense mechanism against

infection (Malacarne et al. 2011). Similarly, for pallidol, evidence exists of its role in resistance of grapevine against *P. viticola* (Pezet et al. 2004). The accumulation of these metabolites in inoculated grapevines could explain the ability of *E. ludwigii* EnVs6 to control *P. viticola* infection on grapevine leaf discs (Campisano et al. 2014c).We propose a possible mechanism by which endophytic colonization might trigger a plant response that leads to the depletion of particular metabolites, enhancing thus the “balanced antagonism” that has been described before for other endophytic organisms (Schulz et al. 1999) In conclusion, two of the three endophytes isolated from grapevine can recolonize grapevine plantlets as endophytes. *E. ludwigii* EnVs6 left a metabolic signature that is characterized by the accumulation of hydroxycinnamic acids and flavonoids and a decrease in phytoalexins. This metabolic signature may reflect the process of colonization and suggests the existence of a possible biological marker associated with endophytism. The mentioned metabolites are expected to counter the penetration of pathogens in grapevine plants when simultaneously accumulated. In our experiments their accumulation is associated to the successful and non-symptomatic colonization of the root endosphere.

5. CONCLUSIONS AND FUTURE WORK

Symbiosis of grapevine and microbes is just beginning to draw attention of farmers, the wine industry and researchers all over the world. One of the main reasons for such interest is the effect that eco-geographical conditions may have on grapevines and therefore wine quality. ***Terroir*** is the effect that biogeographical circumstances impart to wine quality (Gilbert et al., 2014). This initially intuitive definition has been backed up by experimental data suggesting that even slight changes in environmental characteristics of the vineyard can lead to contrastingly different grape and thus wine (Zarraonaindia and Gilbert, 2015). The term has thus been used to describe the relationship between wine´s physicochemical characteristics and the growth conditions in which plants are cultivated, including climate, soil characteristics, fertilization and pest management strategies. Clearly, one of the main factors involved in wine´s *terroir* are microbes, since they are the most abundant living organisms in the environment (David, 2012) and can engage in multiple ecological interactions (Braga et al., 2016). Understanding the role of microbial endophytes in grapevine biology and most importantly elucidating how they (positively or negatively) affect the plant can result in endophyte-based biotechnologies for crop protection. On the other hand this could be translated into the production of **healthier grapes of a higher quality** as a result of plant protection provided by the grapevine symbionts. Moreover, if *terroir* can be linked to the microbiome of the plant, particularly to the symbionts living inside the host, the use of engineered endophytic communities for the modification of wine properties could become a GMO-free approach for better wine production.

The goal of this work was to analyze how bacterial endophytes interact with grapevine and with other grapevine symbionts at the cellular and molecular levels. This unprecedented investigation was motivated by the idea of **using endophytes as agents of plant protection** not only through the classical approach of “one strain, one enzyme”, but also by understanding the ecology of the whole endophytic community and its dynamics in the vineyard. We envisioned endophytes as a rich source of gene functions useful not only in agricultural practices, but also in other instances, such as the pharmaceutical industry, bioremediation, and microbial enhanced oil recovery (MEOR). Most importantly, the rationale behind the study of the whole endophytic community and its biotechnological potential, its transmission and its effect in grapevine plants was based on the premise of using endophytes as a vaccination system in plants. This groundbreaking idea comes from previous observations showing that bacterial communities are involved in disease suppression. The so-called “suppressive soils” where disease seems to be arrested in conditions that are otherwise favorable (Kinkel et al., 2011) are now a paradigm of the effect of microbial communities on plant health. These soils are subject to forces that shape their microbial communities in a way that most antagonists of plant pathogens are enriched. These interact with environmental factors to create a perfect scenario for disease suppression. In human medicine, bacterial communities are also involved in disease suppression. As an example, fecal transplants can dramatically reduce the risk of developing gastrointestinal pathologies and to some extent reverse the effect of pathogenic microorganisms in the gut. It is thus relevant to clarify that the term “vaccine” here does not refer to antigens for acquired immunity, but to the use of endophytes as “**probiotics**” (Fox, 2015).

But beyond the practical considerations, we wanted to contribute to the definition of endophyte by finding **unique characteristics** that represent the endophytic lifestyle and by understanding how

genome and phenotype define these symbionts. Although microorganisms can be isolated virtually from inside all plants on earth (in nature a truly axenic plant very rarely occurs and has an adaptive disadvantage; Partida-Martínez and Heil, 2011) recognizing who is an endophyte and who is not is not trivial. The definition of endophyte is in itself difficult to interpret. Are all organisms living inside plants endophytes? If so, can we consider acute infections in plants as endophytes? What about microbial contaminations? What happens to organisms that reside in the plant, apparently asymptomatically, but under certain conditions express virulence factors and become detrimental? Are these also endophytes? It becomes clear that It is not an easy task to differentiate endophytes from non-endophytic organisms.

The results of this work suggest that symbiosis of endophytes with grapevine is dynamic. We have shown that there are genetic elements shared with pathogens and epiphytes that populate endophytic genomes (see Figures 6-8), proving that endophytism is associated with multiple lineages and plausibly multiple origins, i.e. the endophytic lifestyle has been inherited from multiple ancestors, but could have disappeared and been regained in some taxa during the course of their evolution, as pointed out elsewhere (Xu et al., 2014). The fact that genomes of endophytic bacteria isolated from grapevine resemble phylogenetically related plant and human pathogens as well as epiphytes might also mean that the endophytic lifestyle could have appeared as a consequence of **niche segregation** (Freilich et al., 2011; Sheppard et al., 2011). With our data, we cannot conclude with precision if these genomic traits shared among our endophytic isolates and epiphytes, pathogens and other endophytes are involved in the niche segregation phenomenon, especially since our number of samples is limited. It is certain nonetheless that these genetic functions are differentially present in endophytes and differentiate them from the other lifestyles. For example, in Figure 7 we exemplified how the accessory genomes of endophytic strains have functions that are not present in non-endophytic strains; in some cases genomes of endophytes from grapevine look more similar to genomes of endophytes from other hosts than to those of non-endophytes.

There is no consensus regarding the origin of endophytes. Some authors have proposed that endophytes may be “disarmed pathogens” that have evolved from a virulent ancestry (Reinhold-Hurek and Hurek, 2011), while others consider endophytism to be an archaic form of non-pathogenic microorganism that remained trapped in the plant under adaptive constraints. Our data however suggest that other constraints could have played roles in endophytic evolution and consequently adaptation to this particular ecological niche. The strain *Erwinia* ErVv1 and the *Pantoea* strains isolated from grapevine recall the epiphytic lifestyle. In contrast, in grapevine’s endophytic *Enterobacter* strains, genomic structure is more similar to that of the saprobiotic or parasitic lifestyle. In fact, our *Enterobacter* strains possess many virulence factors that can act as mechanisms for chronic infection. This is the case with the toxin-antitoxin systems that have been implicated in high virulence of *Mycobacterium tuberculosis*, which provoke long dormancy and persistence inside the host (Yamaguchi et al., 2011). These data support the hypothesis that epiphytic, saprophytic, saprobiotic and pathogenic lifestyles in grapevine could have been derived in a very specialized endophytic niche to live in intimate relationship with the plant, and to efficiently use the host resources. In that sense, the endophytic niche might have arisen as a specialization of those lifestyles during co-evolution. It seems that niche occupation by endophytes of grapevine is achieved by using the same genetic

arsenal present in epiphytes or pathogens (see Appendix Table 9, where secretion systems, well known virulence factors in pathogenic bacteria, can also be present in endophytes and even in epiphytes), but possibly employed in a different fashion. This variability might be related to several selective pressures in the vineyard, including domestication, competition in the agro-ecosystem and horizontal gene transfer, under which some genomic traits might be fixed.

Domestication as one of the most important events of grapevine's evolution has had a tremendous impact on the plant's microbiome (Pinto et al., 2014). It cannot be excluded that, as was the case for *P. acnes* type Zappae (Campisano et al., 2014), niche specialization of parasites or free-living organisms evolving into a mutualistic lifestyle came hand in hand with domestication of vine plants. A strong body of evidence supports this statement, by localizing the phylogenies of endophytic fungi and hosts plants together in a co-cladogenesis phenomenon (Clay and Schardl, 2002). Evidence for the involvement of the domestication process in the evolution of endophytic genomes of grapevine is the presence of sensor and detoxifying proteins for heavy metals. In the genomes of our *Enterobacter* and *Pantoea* strains, we found genes coding for copper efflux pumps as well as genes related to detoxification of Cd^{2+}, Zn^{2+} and Co^{2+}.(Appendix Table 5 and 7). These genes are uniquely present in some of our strains and not in other free living species. These genes represent adaptations of the microbes to conditions like amendment of soil for fertilization and antifungal treatments (Turner et al., 2013).

Competition is another important selector that might be reflected in the genomes of endophytes. Some of our grapevine isolates possess a plethora of genes dedicated to counter competitors. These include antibiotic synthesis and two component systems (similar to the *evgA or* the *cre* operon of *E. coli)* that confer resistance to antibiotics (see Appendix Table 5-7). To outgrow and protect themselves from many bacterial and fungal competitors, endophytes most endure a chemical warfare in the endobiome. These events must have shaped the genomes of endophytes to such an extent that even some gene functions have been fixed as a sign of endophytism, sometimes being present but not expressed. But in other cases, the genes might be absent due to a lack of use or exposure to antibiotic producing antagonists. Our data confirm that some of the genes for antibiotic synthesis are actually expressed (for example antibiotic resistance tests in Appendix Figure 1 show that bacterial endophytes actually express antibiotic resistance genes, found in their genomes as shown in Appendix Tables 5-7) and that their expression might be modulated by ecological interaction with other endophytes occupying the same niche (Figure 28).

Horizontal gene transfer might also be implicated in the shaping of endophytic genomes. This research shows that endophytes from grapevine have a large number of transposable elements (Appendix Tables 5-8). Also, data comparing genome sizes among *Pantoea* strains show that PaVv9 has the largest genome pf the *Pantoea* analyzed here, consisting of 9,754,510 base pairs (In Appendix Table 2 the genome size of strain PaVv9 is also reflected in the number of contigs assembled - 385 – the largest from all endophytes sequenced). We also have provided evidence for the presence of phage genes (Apendix Table 8) inside the genomes of the endophytes as well as CRISPR sequences (Table 4) that play an important role in allelic diversity in bacteria. These episomes, as well as mobile and phage elements might be indicators of horizontal flux that explain why so many endophytes contain gene functions that apparently would not have a value in their hosts (collagen degrading enzymes for example). While some endophytic strains may use these elements to increase fitness

due to exchange of advantageous functions, some of them might have them as traces of an ancient life where severe environmental conditions dominated and gene exchange was important for survival.

We have also shown that endophytes elicit a metabolic response in the plant (Figures 17), which is associated with flavonoid regulation (Appendix Table 13) and possibly with control of the synthesis of phytoalexins (substances that protect the plant from pathogen attack). This metabolic signature may well have evolved as a strategy for optimal and selective colonization of endophytes because depletion of some of these metabolites (for example *trans*-resveratrol, and α- and ε-viniferin) have been shown to represent a decrease of function in the first line of self-defense in the plant (Ahhuja et al., 2012). These phytoalexins were found differentially accumulated between roots and stems of inoculated plants (Figure 20) being less concentrated in the latter. Compartmentalization of metabolic responses in plants might be a consequence of plant morphology, but it may also be a systemic effect of grapevine colonization by endophytes. The responses may not just be local, but may extend to other plant compartments, ensuring the effective spreading of the colonizing microbe. We propose that further phylogenetic studies should be performed with a larger set of samples, using multiple molecular markers to evaluate the distribution of endophytic bacteria from grapevine in the natural history of endophytic and non-endophytic microorganisms.

We have proven that the entire bacterial endophytic community can be transferred between plants through *S. titanus* (Figure 9). This reflects the holobiont status of grapevine plants that have co-evolved with insects and microorganisms (and possibly many other organisms).
If studied in more detail, all of the findings above, i.e. the vector assisted transmission of endophytic bacteria, their enzymatic potential for plant protection, the knowledge of their genomic structure and organization together with the metabolic effect on plant hosts icould be translated as means of a natural exploitation of interactions for effective use of the microbiota in plant.

5.1. Markers for endophytism

We looked for candidate processes or mechanisms that could eventually be used for defining endophytes. The isolation of microorganisms from surface-sterilized plants is not sufficient to define the endophytic lifestyle of an isolate, because the definition presented here is based on ecological constraints rather than technicalities.

To find such candidate markers, we performed comparisons based on orthology, defining orthologue as a gene derived from a common ancestor and replicated through speciation. This allowed us to propose that the structural similarities (gene content, GC content, size of core and accessory genomes) observed in any given pair of genomes in our work represent characteristics that have been inherited from a common ancestor and so they truly reflect traits conserved in the endophytic lifestyle. Moreover, unique genomic functions, shared between endophytes from grapevine and from other hosts (Figure 2) might be characteristic of endophytic genomes, since they were not detected in bacteria with other lifestyles (Appendix Tables 5-8 and Table 4).

These similarities and unique characteristics of endophytes in contrast to non-endophytes should be further explored to be considered as **first markers** for defining endophytism. A precautious interpretation of these findings must be considered, however, since: a) a bias might have been introduced by threshold values for best reciprocal hit comparisons and other metrics and b) whether or not those shared characteristics are expressed has to be experimentally determined. Thus, these observations should be validated with a larger set of data and more detailed experiments involving transcription profiles and genomic libraries for confirmation of functions.

We have identified a change in plant secondary metabolism that is imprinted when endophytes colonize the host (Figures 18-19) and that can be considered a “metabolic signature” (Scherling et al., 2009). These changes could be reproduced in independent experiments in which grapevine plantlets were inoculated with *Enterobacter ludwigii* EnVs6 (Figure 20). It is not for sure that this phenomenon is reproducible in other holobionts, though our data shows a consistent shift in grapevine metabolism due to endophytic bacterial colonization (Figures 18-19), which is absent in plants inoculated with non-endophytic bacteria (Figure 18-19). This is evidence for a possible **second marker** for the definition of endophytism, but further investigations need to be performed using both bacterial and plant mutants to confirm that the metabolic response in the plant is solely due to colonization of bacteria and not to other factors.

Other observations in our experiments show non-recurrent mechanisms that do not seem to be target markers for endophytism. The colonization experiments using *in situ* hybridization have shown how metabolically active endophytes enter the plantlets through secondary roots, as has been reported in other grapevine endophytes colonizing through root hairs and branching, and are able to colonize the plant interior in a very short time (Figures 14-16). Multiple pathogens and other symbionts (Rhizobia and *Frankia* diazotrophs, for example) can use the same colonization routes. Thus, it is plausible that these points of entry might be selected just because they are physically accessible to microorganisms. This means of infecting the host does not contradict the fact that plant exudates and other chemical signals may be involved in the attraction of those microorganisms to the host.

Our colonization experiments also showed that *Sphingomonas phyllosphaere* SpVs6 isolated from grapevine was not able to colonize the plant in an *in vitro* setting (Figure 16). This is of great relevance, since it reflects the ambiguity of the definition of endophyte. We have questioned if the conditions provided in our experiments were sufficient to meet the requirements for endophytism (Lopez-Fernandez et al., 2015). Although conditions were used that were favorable for growth of the host and with media composed of nutrients found in nature, it is possible that endophytism depends also on the interactions and chemical signaling within the whole community. If that is true, *Sphingomonas phyllosphaere* SpVs6 could also be a facultative endophyte that depends of such factors to behave as an endophyte. In the rhizosphere of plants, gene functions for quorum sensing and chemical communication have been identified in other holobionts (Schaefer et al., 2013). These genes have special features in recognizing plant-derived compounds.

Although bacterial colonization was apparently beneficial for the plant (high production of IAA and ACC, as well as antibiotic resistance, see Appendix Figure 1 and Figures 3 and 4), our genomic data

coupled with phenotypical attributes of the bacterial endophytes highlights how such beneficial properties could hypothetically be expressed both as virulence factors and as plant protection molecules. Hence, the activity of these molecules could also be an advantage for non-symbionts or even for pathogens. Certainly, the production of IAA, QS molecules, ACC and enzymes can equally benefit endophytic symbionts and pathogens to maintain their association with the plant and to acquire nutrients necessary for their replication.

The transmission of a bacterial endophytic community of grapevine also harbors elements that could exemplify a unique lifestyle. Experiments of transmission through *S. titanus* showed that a whole community of endophytes can be transported between grapevine plants (Figure 9) and that community composition during transmission was affected by the host. Data presented here shows that some of the transferred bacterial taxa seemed to cope more easily with host switching (Figures 11-12). Additionally, some were efficiently transferred, while others were not acquired by the insect (Figure 13). This also suggests that within the whole community not all members are active endophytes. Putting all these observations together, we theorize that sub-populations in the plant-dwelling community may arise, containing organisms with two niches.

According to our observations, plant-dwelling bacteria can live associated with the host and engage in interactions with the host. These **symbiotic endophytes** (for example in Appendix Figure 1 the best performance in plant protection properties represented by a black total score in last column) are the ones that have evolved a genomic and phenotypic arsenal for association with the plant and from our perspective, should be the ones targeted as potential beneficial organisms (here, as defined by Zook (Zook, 2015),symbiosis is the acquisition of one organism by another unlike organism and, resulting from subsequent long time intimacy, new structures and metabolisms develop)

All the same, there are members of these plant dwelling bacteria (found also in our experiments) that are not active and live in the host in a neutral way. (See, for example in Appendix Figure 1, the worst performance in plant protection properties represented by a white total score in last column. Compare also the colonization efficiency of *Sphingomonas phyllosphaere* SpVs6 to that of *Enterobacter ludwigii* EnVs6 in Figures 14 and 16 and the transmission of *Enterobacter ludwigii* EnVs6 with that of *Pantoea vagans* PaVv9 in Figure 13). These **non-symbiotic endophytes** live passively inside the plant without causing any harm but not engaging in active ecological and metabolic interactions. Considering our observations, we propose that the symbiotic endophytes meet the (non-excluding) criteria defined here as markers. It is important to highlight that our data also suggest symbiotic endophytism is not a fixed and constant state in an organism's life cycle as has been previously suggested (Schulz and Boyle, 2005). Whether or not endophytism is mutualistic for the partners depends on the interactions between various factors: a) the interaction between host defense and endophytic virulence factors, i.e. a balance of antagonisms as hypothesized by Schulz et al. (1999) for fungal endophytes, b) interactions with other members of the community, and c) physical and chemical factors and as we have shown, and d) the virulence potential encoded in their genomes.

With the variety of gene functions that code for beneficial and detrimental molecules and the regulatory mechanisms of gene expression, each endophyte has the potential to switch lifestyles at any given point. We believe this is in agreement with numerous observations showing that the

beneficial properties of bacteria and the virulence of pathogens are separated by a thin layer of interactions.

5.2. Use of endophytes as a prophylactic tool in grapevine plants

Besides thinking about characteristics for the definition of endophytes in grapevine, the experiments presented in this thesis provide **evidence** for the potential use of endophytes as a natural vaccination (in the sense of this thesis, see 1.1) strategy in agreement with previous research showing that microbial communities applied to plants can function as a probiotic strategy for plant health (Spence and Bais, 2013). Immunity of plants is a finely regulated system that enables recognition of self from non-self. And as has been clearly exposed in many investigations, immunity can also be primed in plants by exposing individuals to antigens that are present in microbes, although without an acquired immunity (Dutta et al., 2014).

As we have shown, the beneficial properties of an endophytic community (as a composite) are not only influenced by genome organization (Appendix Tables 5-7), but also by the host it inhabits (Figure 3 and Appendix Figure 1). Therefore, formulations, i.e. consortia of endophytes applied in the field either as sprays, amendments or other delivery system, must be developed and designed as individualized therapies rather than as a generic “one size fits all” strategy. Using epigenetic manipulation of endophytic communities to take advantage of the key players in symbiosis, such formulations could become less toxic and become more efficient in plant protection. An accurate identification of truly mutualistic endophytes that are metabolically active, able to prime the immune system of the plant and maintain neutral to beneficial interactions with other members of the community is then a crucial step when planning endophyte based therapies. Also, evaluating the virulence potential of endophytic isolates could help in the selection of avirulent strains for use in crops Surely, there is no standard or ideal endophytic community that can be beneficial for all plants. Rather, the beneficial or detrimental potential relies on multiple ecological interactions in the endo -, rhizo- and phyllosphere.

We have shown that individual endophytes can also display potent beneficial properties. An example of one such symbiont is our strain *Enterobacter ludwigii* EnVs6 that was capable of effectively colonizing the plant (Figure 14), imprinting a metabolic signature (Figure 18-20), inducing growth to inoculated grapevine plants and protecting them from fungal pathogens (Appendix Figure 1), while being efficiently transferred from source to sink plants through *S. titanus* (Figure 13A). Nonetheless, this partially correlates with the true nature of the endophyte in the community since the mode of action of plant beneficial properties might be regulated at the community level. It is therefore reasonable to think that PGP and biocontrol properties of the microbial community will surface as an emergent property rather than the sum of individual enzymatic activities and beneficial potentials.

Our experiments provide evidence for the horizontal transmission of the endophytic community of grapevine by *S. titanus*. This to our knowledge is a pioneer work and a milestone in the development of a natural delivery system for beneficial microbes in grapevine crops. Not only because we have proven that the endophytic community is efficiently (Figure 10) and selectively (Figure 13) transferred by a grapevine-associated insect, but because we have demonstrated the enzymatic diversity of endophytic communities and its variation across hosts (Figure 3, Figure 11, and Appendix Figure 2-3). Those observations put endophytes in a status of important agents of crop management, whose

bioactivities not only work locally, but could also be transferred. This should prompt researchers to think in more holistic approaches to define consortia of beneficial bacteria for their use in crop protection and to change the paradigm from selectively isolated beneficial microbes to one that employs and exploits microbial communities and their interactions to obtain a stably healthy crop.

The phenotypical capabilities that we have shown such as plant growth promotion, biocontrol and antibiosis (Appendix Figure 1; Figures 4-5) were also confirmed at the genomic level (Appendix Tables 5-7). The correlation of phenotypical and genomic findings indicates that the symbiotic endophytic lifestyle is an active metabolic state in which many genes are selectively transcribed. Heterogeneity of the endophytic community and its relation with enzymatic potential also suggests that functional subgroups, which are linked to taxonomical groups (Appendix Figure 3), may play a role in symbiosis with the plant. As presented in this work, there are members of the endophytic community whose activities remarkably affect plant homeostasis, while others, even though isolated from the same plant host, remain inactive in a cryptic way.

A model is proposed for interactions in the grapevine holobiont, in which endophytic bacteria can be acquired by the insect (Figure 21, Step 1) and be selectively transferred (Figure 21, Step 2) from plant to plant. During the transition process, the vector shapes the community structure that finally arrives at the plant (Figure 21, Step 3). Colonization occurs when active endophytes effectively overcome the plant's defense, including their constitutive and induced secondary metabolites (Figure 21, Step 4), infecting the plant through the stems. Other endophytes colonize through the roots crossing barriers in the secondary root branches. Endophytes delivered to the stems grow into the roots where they may also exert a metabolic signature of colonization. Plant metabolites are involved in limiting microbial colonization, so that the bacteria and fungi do not become pathogens. Microbial metabolites have multiple functions, including limiting the growth of the other inhabitants of the community. This cycle is also plausibly regulated by insect physiology (developmental stages and diapause), which in turn are finely tuned with plant physiology. The cycle could repeat itself over and over again maintaining a stable community of microbes *in planta.*

Research in the field of plant probiotics should move from improving strains to improving communities and therefore exploiting ecological interactions rather than only enzymatic aspects of strains.

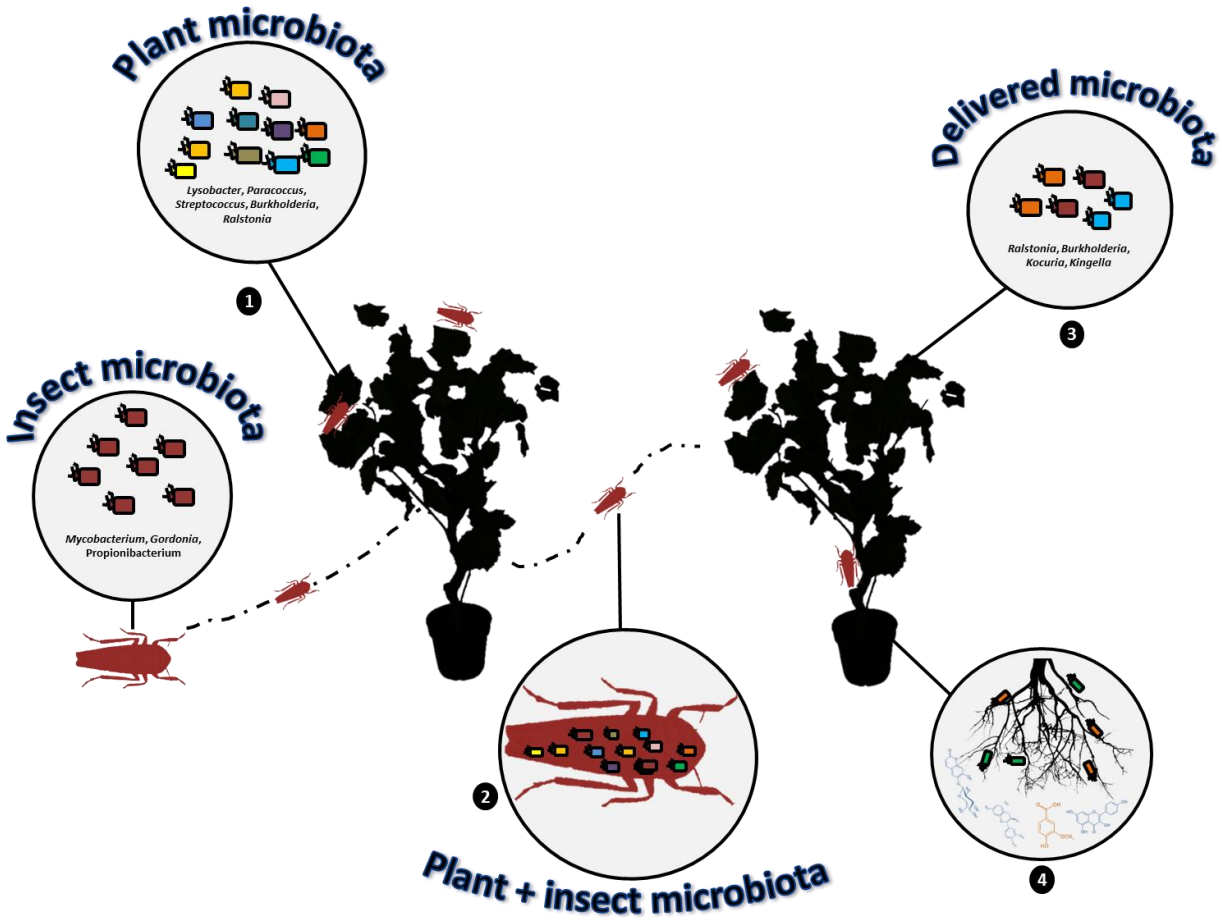


Figure 21. Model of interactions in the **grapevine holobiont**: Plant and insect microbiota are dominated by specific taxa. Plant microbiota is acquired by the insect **(1).** The microbiota of both insect and plant mix and selection within the insect occurs while the insect transfers the microbiota **(2)**. Endophytic microbiota is delivered by the insect to the stems and infects the plant. Microbiota of the insect is dominated by Actinobacteria. Microbiota of the plant is dominated by Proteobacteria. In the vector, both groups can be found. In the delivered microbiota, only a selected group of the original plant microbiota will colonize the plant. **(3)**. Endophytes engage in metabolic interactions with the plant **(4)**.Vanillic acid, kampferol, catechin and other flavonoids are involved in plant defense and as proposed in this work, are depleted upon endophytic colonization.

6. BIBLIOGRAPHY

Abu Kwaik, Y., and Bumann, D. (2013). Microbial quest for food in vivo: 'nutritional virulence' as an emerging paradigm. Cell Microbiol. 15, 882-890. doi: 10.1111/cmi.12138.

Adame-Álvarez, R-M., Mendiola-Soto, J., and Heil, M. (2014). Order of arrival shifts endophyte-pathogen interactions in bean from resistance induction to disease facilitation. FEMS Microbiol Lett. 355, 100–107. doi:10.1111/1574-6968.12454.

Ahmad, N., Hamayun, M., Khan, S. A., Khan, A. L., Lee, I.-J., and Shin, D.-H. (2010). Gibberellin-producing endophytic fungi isolated from *Monochoria vaginalis*. J Microbiol Biotechnol 20, 1744–1749.

Ahmed, N. (2009). A flood of microbial genomes-do we need more? PLoS ONE 4. e5831. doi: 10.1371/journal.pone.0005831.

Akinsanya, M. A., Goh, J. K., Lim, S. P., and Ting, A. S. Y. (2015). Metagenomics study of endophytic bacteria in *Aloe vera* using next-generation technology. Genom Data 6, 159–163. doi:10.1016/j.gdata.2015.09.004.

Alexey, G., Vladislav, S., Nikolay, V., and Glenn, T. (2013) QUAST: quality assessment tool for genome assemblies, Bioinformatics 29, 1072-1075. doi: 10.1093/bioinformatics/btt086

Alikhan, N.F., Petty, N.K., Ben Zakour, N.L., and Beatson, S.A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics. 12, 402. doi: 10.1186/1471-2164-12-402.

Allwood J.W., Ellis, D.I., Goodacre, R. (2008) Metabolomic technologies and their application to the study of plants and plant-host interactions. Physiol plantarum 132, 117-135. doi: 10.1111/j.1399-3054.2007.01001.x.

Amadou, C., Pascal, G., Mangenot, S., Glew, M., Bontemps, C., Capela, D., et al. (2008). Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. Genome Res. 18, 1472 - 1483.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux R, Stahl, D.A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol. 56, 1919-1925.

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32-46.

Atangana, A., Khasa, D., Chang, S., and Degrande, A. (2014). “Ecological Interactions and Productivity in Agroforestry Systems,” in Tropical Agroforestry (Dordrecht: Springer Netherlands), 151–172. Available at: http://link.springer.com/10.1007/978-94-007-7723-1_7 [Accessed January 17, 2017].

Bacilio-Jiménez, M., Aguilar-Flores, S., Ventura-Zapata, E., Pérez-Campos, E., Bouquelet, S., Zenteno, E. (2003) Chemical characterization of root exudates from rice (*Oryza sativa*) and their effects on the chemotactic response of endophytic bacteria. Plant Soil. 249, 271-277. doi: 10.1023/A:1022888900465.

Backert, S., Schwarz, T., Miehlke, S., Kirsch, C., Sommer, C., Kwok, T., et al. (2004). Functional analysis of the cag pathogenicity island in *Helicobacter pylori* isolates from patients with gastritis, peptic ulcer, and gastric cancer. Infect Immun. 72, 1043-1056.

Balachandar, D., Sandhiya, G. S., Sugitha, T. C. K., and Kumar, K. (2006). Flavonoids and growth hormones influence endophytic colonization and *in planta* nitrogen fixation by a diazotrophic *Serratia* sp. in rice. World J Microbiol Biotechnol. 22, 707–712. doi:10.1007/s11274-005-9094-0.

Barelli, L., Moonjely, S., Behie, S. W., and Bidochka, M. J. (2016). Fungi with multifunctional lifestyles: endophytic insect pathogenic fungi. Plant Mol Biol. 90, 657–664. doi:10.1007/s11103-015-0413-z.

Bennett, A. E. (2013). Can plant-microbe-insect interactions enhance or inhibit the spread of invasive species? Funct.Ecol 27, 661–671. doi:10.1111/1365-2435.12099.

Betz, F. S., Hammond, B. G., and Fuchs, R. L. (2000). Safety and advantages of *Bacillus thuringiensis*-protected plants to control insect pests. Regul Toxicol Pharmacol. 32, 156–173. doi:10.1006/rtph.2000.1426.

Bhore, S., Christina, A., and Christapher, V. (2013). Endophytic bacteria as a source of novel antibiotics: An overview. Pharmacogn Rev. 7, 11. doi:10.4103/0973-7847.112833.

Biagini, B., G.D. Lorenzis, A. Scienza, O. Failla, S. Imazio, and D. Maghradze. (2012). Wild grapevine (*Vitis vinifera* L. subsp. *sylvestris* (Gmelin) Hegi) in Italy: distribution and preliminary genetic analysis. Acta hort. 948, 211 - 216.

Binnewies, T. T., Motro, Y., Hallin, P. F., Lund, O., Dunn, D., La, T., et al. (2006). Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. Funct Integr Genomics 6, 165–185. doi:10.1007/s10142-006-0027-2.

Bland, C., Ramsey, T.L., Sabree, F., Lowe, M., Brown, K., Kyrpides, N.C., et al. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 8, 209. doi: 10.1186/1471-2105-8-209.

Bloemberg GV, Wijfjes AH, Lamers GE, Stuurman N, Lugtenberg BJ (2000) Simultaneous imaging of *Pseudomonas fluorescens* WCS365 populations expressing three different autofluorescent proteins in the rhizosphere: new perspectives for studying microbial communities. Mol Plant Microbe Interact. 13, 1170-1176. doi: 10.1094/mpmi.2000.13.11.1170.

Bordenstein, S. R., and Theis, K. R. (2015). Host biology in light of the microbiome: ten principles of holobionts and hologenomes. PLOS Biol. 13, e1002226. doi:10.1371/journal.pbio.1002226.

Brader, G., Compant, S., Mitter, B., Trognitz, F., and Sessitsch, A. (2014). Metabolic potential of endophytic bacteria. Curr Opin Biotechnol. 27, 30–37. doi:10.1016/j.copbio.2013.09.012.

Braga, R. M., Dourado, M. N., and Araújo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. Braz. J Microbiol. 47, 86–98. doi:10.1016/j.bjm.2016.10.005.

Braun, K., Romero, J., Liddell, C., and Creamer, R. (2003). Production of swainsonine by fungal endophytes of locoweed. Mycol Res. 107, 980–988. doi:10.1017/S095375620300813X.

Brenner, D.J., Mcwhorter, A.C., Kai, A., Steigerwalt, A.G., and Farmer, J.J., 3rd (1986). *Enterobacter asburiae* sp. nov., a new species found in clinical specimens, and reassignment of *Erwinia dissolvens* and *Erwinia nimipressuralis* to the genus *Enterobacter* as *Enterobacter dissolvens* comb. nov. and *Enterobacter nimipressuralis* comb. nov. J Clin Microbiol. 23, 1114-1120.

Brunori, E., Farina, R., and Biasi, R. (2016). Sustainable viticulture: the carbon-sink function of the vineyard agro-ecosystem. Agric Ecosyst Environ 223, 10–21. doi:10.1016/j.agee.2016.02.012.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L., and Schulze-Lefert, P. (2013). Structure and functions of the bacterial microbiota of plants. Annu Rev Plant Biol. 64, 807–838. doi:10.1146/annurev-arplant-050312-120106.

Bulgari, D., Casati, P., and Faoro, F. (2011). Fluorescence *in situ* hybridization for phytoplasma and endophytic bacteria localization in plant tissues. J Microbiol Methods. 87, 220–223. doi:10.1016/j.mimet.2011.08.001.

Bulgari, D., Casati, P., Brusetti, L., Quaglino, F., Brasca, M., Daffonchio, D., et al. (2009). Endophytic bacterial diversity in grapevine (*Vitis vinifera* L.) leaves described by 16S rRNA gene sequence analysis and length heterogeneity-PCR. J Microbiol. 47, 393–401. doi:10.1007/s12275-009-0082-1.

Burdman S, Jurkevitch E, Schwartsburd B, Okon Y (1999) Involvement of outer-membrane proteins in the aggregation of Azospirillum brasilense. Microbiology. 145 (Pt 5): 1145-1152.

Campisano, A., Antonielli, L., Pancher, M., Yousaf, S., Pindo, M., Pertot, I. (2014a). Bacterial Endophytic Communities in the Grapevine Depend on Pest Management. PLoS ONE 9, e112763. doi: 10.1371/journal.pone.0112763.

Campisano, A., Ometto, L., Compant, S., Pancher, M., Antonielli, L., Yousaf, S., et al. (2014b). Interkingdom transfer of the acne-causing agent, *Propionibacterium acnes*, from human to grapevine. Mol Biol Evol. 31, 1059–1065. doi:10.1093/molbev/msu075.

Campisano, A., Pancher, M., Puopolo, G., Puddu, A., Lòpez-Fernàndez, S., Biagini, B., et al. (2014c). Diversity in endophytic populations reveals functional and taxonomic diversity between wild and domesticated grapevines. Am J Enol Vitic. doi: 10.5344/ajev.2014.14046.

Capoen, W., Den Herder, J., Sun, J., Verplancke, C., De Keyser, A., De Rycke, R., et al. (2009). Calcium spiking patterns and the role of the calcium/calmodulin-dependent kinase CCaMK in lateral root base nodulation of *Sesbania rostrata*. Plant cell. 21, 1526-1540. doi: 10.1105/tpc.109.066233.

Caporaso, J. G., Bittinger, K., Bushman, F. D., DeSantis, T. Z., Andersen, G. L., and Knight, R. (2010a). PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics. 26, 266–267. doi:10.1093/bioinformatics/btp636.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010b). QIIME allows analysis of high-throughput community sequencing data. Nat Methods. 7, 335–336. doi:10.1038/nmeth.f.303.

Carrell, A. A., and Frank, A. C. (2015). Bacterial endophyte communities in the foliage of coast redwood and giant sequoia. Front Microbiol. 6. doi:10.3389/fmicb.2015.01008.

Cha, C., P. Gao, Y.C. Chen, P.D. Shaw, and S.K. Farrand. (1998). Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. Mol Plant Microbe Interact. 11, 1119-1129.

Chaturvedi, H., and Singh, V. (2016). Potential of bacterial endophytes as plant growth promoting factors. J Plant Pathol Microbiol. 7. doi:10.4172/2157-7471.1000376.

Chaudhry, V., and Patil, P. B. (2016). Genomic investigation reveals evolution and lifestyle adaptation of endophytic *Staphylococcus epidermidis*. Sci Rep. 6. doi:10.1038/srep19263.

Chen, L., Xiong, Z., Sun, L., Yang, J., and Jin, Q. (2012). VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic Acids Res. 40, D641-645. doi: 10.1093/nar/gkr989.

Choi, O., Lim, J.Y., Seo, Y.-S., Hwang, I., and Kim, J. (2012). Complete genome sequence of the rice pathogen *Pantoea ananatis* strain PA13. J Bacteriol. 194, 531-531.

Chuche, J., and Thiéry, D. (2014). Biology and ecology of the Flavescence dorée vector Scaphoideus titanus: a review. Agron Sustain Dev. 34, 381–403. doi:10.1007/s13593-014-0208-7.

Clarke, K.R. (1993). Nonparametric multivariate analyses of changes in community structure. Austral J Ecol. 18, 117-143.

Clay, K. (1993). The ecology and evolution of endophytes. Agric Ecosyst Environ. 44, 39–64. doi:10.1016/0167-8809(93)90038-Q.

Clay, K., and Schardl, C. (2002). Evolutionary origins and ecological consequences of endophyte symbiosis with grasses. Am Nat. 160 Suppl 4, S99-S127. doi: 10.1086/342161.

Compant., S., B. Reiter, A. Sessitsch, J. Nowak, C. Clement, and Barka, E.A. (2005). Endophytic colonization of *Vitis vinifera* L. by plant growth promoting bacterium *Burkholderia* sp. strain PsJN. Appl Environ Microbiol. 71, 1685-1693.

Compant, S., Clement, C., and Sessitsch, A. (2010). Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem. 42, 669-678.

Compant, S., Kaplan, H., Sessitsch, A., Nowak, J., Ait Barka, E., and Clément, C. (2008). Endophytic colonization of *Vitis vinifera* L. by *Burkholderia phytofirmans* strain PsJN: from the rhizosphere to inflorescence tissues. FEMS Microbiol. Ecol. 63, 84–93. doi:10.1111/j.1574-6941.2007.00410.x.

Compant, S., Mitter, B., Colli-Mull, J. G., Gangl, H., and Sessitsch, A. (2011). Endophytes of grapevine flowers, berries, and seeds: identification of cultivable bacteria, comparison with other plant parts, and visualization of niches of colonization. Microb Ecol. 62, 188–197. doi:10.1007/s00248-011-9883-y.

Daims, H., Bruhl, A., Amann, R., Schleifer, K-H, Wagner, M. (1999). The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol. 22, 434-444. doi: 10.1016/s0723-2020(99)80053-8.

Darling, A.C., Mau, B., Blattner, F.R., and Perna, N.T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 14, 1394-1403. doi: 10.1101/gr.2289704.

David, R. (2012). Environmental microbiology: re-evaluating the abundance of microorganisms. Nat Rev Microbiol. 10, 671–671. doi:10.1038/nrmicro2889.

De Maayer, P., Chan, W.Y., Venter, S.N., Toth, I.K., Birch, P.R., Joubert, F., et al. (2010). Genome sequence of *Pantoea ananatis* LMG20103, the causative agent of *Eucalyptus* blight and dieback. J Bacteriol. 192, 2936-2937. doi: 10.1128/jb.00060-10.

Decante, D., and van Helden, M. (2006). Population ecology of *Empoasca vitis* (Göthe) and *Scaphoideus titanus* (Ball) in Bordeaux vineyards: influence of migration and landscape. Crop Prot. 25, 696–704. doi:10.1016/j.cropro.2005.09.016.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., et al. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ. Microbiol. 72, 5069–5072. doi:10.1128/AEM.03006-05.

Diao, H., Yan, H. L., Xiao, Y., Yu, B., Yu, J., He, J., et al. (2016). Intestinal microbiota could transfer host gut characteristics from pigs to mice. BMC Microbiol. 16, 238 doi:10.1186/s12866-016-0851-z.

Dutta, D., Puzari, K. C., Gogoi, R., and Dutta, P. (2014). Endophytes: exploitation as a tool in plant protection. Braz Arch Biol. Technol. 57, 621–629. doi:10.1590/S1516-8913201402043.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461. doi:10.1093/bioinformatics/btq461.

Elasri, M., Delorme, S., Lemanceau, P., Stewart, G., Laue, B., Glickmann, E., et al. (2001). Acyl-homoserine lactone production is more common among plant-associated *Pseudomonas* spp. than among soilborne *Pseudomonas* spp. Appl Environ Microbiol. 67, 1198–1209. doi:10.1128/AEM.67.3.1198-1209.2001.

Espinosa-Urgel, M., Kolter, R., and Ramos, J.-L. (2002). Root colonization by *Pseudomonas putida*: love at first sight. Microbiology 148, 341–343. doi:10.1099/00221287-148-2-341.

Euzéby, J.P. 1997. List of bacterial names with standing in nomenclature: a folder available on the Internet. Int J Syst Bacteriol. 47, 590-592. 10.1099/00207713-47-2-590

Filloux, A., Hachani, A., and Bleves, S. (2008). The bacterial type VI secretion machine: yet another player for protein transport across membranes. Microbiology 154, 1570-1583. doi: 10.1099/mic.0.2008/016840-0.

Fisher, P. J., and Petrini, O. (1992). Fungal saprobes and pathogens as endophytes of rice (Oryza sativa L.). New Phytol. 120, 137–143. doi:10.1111/j.1469-8137.1992.tb01066.x.

Fornasiero, D., Pavan, F., Pozzebon, A., Picotti, P., and Duso, C. (2016). Relative infestation level and sensitivity of grapevine cultivars to the leafhopper *Empoasca vitis* (Hemiptera: Cicadellidae). J Econ Entomol. 109, 416–425. doi:10.1093/jee/tov313.

Fox, J. L. (2015). Agricultural probiotics enter spotlight. Nat Biotechnol. 33, 122–122. doi:10.1038/nbt0215-122.

Freilich, S., Zarecki, R., Eilam, O., Segal, E. S., Henry, C. S., Kupiec, M., et al. (2011). Competitive and cooperative metabolic interactions in bacterial communities. Nat Commun. 2, 589. doi:10.1038/ncomms1597.

Fujita, Y., and Fujita, T. (1987). The gluconate operon *gnt* of *Bacillus subtilis* encodes its own transcriptional negative regulator. Proc Natl Acad Sci. U.S.A. 84, 4524-4528.

Gabaldón, T., and Koonin, E. V. (2013). Functional and evolutionary implications of gene orthology. Nat. Rev. Genet. 14, 360–366. doi:10.1038/nrg3456.

Gaiero, J. R., McCall, C. A., Thompson, K. A., Day, N. J., Best, A. S., and Dunfield, K. E. (2013). Inside the root microbiome: bacterial root endophytes and plant growth promotion. Am J Bot. 100, 1738–1750. doi:10.3732/ajb.1200572.

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the “Gene-Jockeying” tool. Microbiol Mol Biol Rev. 67, 16–37. doi:10.1128/MMBR.67.1.16-37.2003.

Ghodsi, M., Hill, C.M., Astrovskaya, I., Lin, H., Sommer, D.D., Koren, S., et al. (2013). *De novo* likelihood-based measures for comparing genome assemblies. BMC Res Notes. 6, 334. doi: 10.1186/1756-0500-6-334.

Gilbert, J. A., van der Lelie, D., and Zarraonaindia, I. (2014). Microbial terroir for wine grapes. Proc Natl Acad Sci. 111, 5–6. doi:10.1073/pnas.1320471110.

Gonella, E., Pajoro, M., Marzorati, M., Crotti, E., Mandrioli, M., Pontini, M., et al. (2015). Plant-mediated interspecific horizontal transmission of an intracellular symbiont in insects. Sci Rep. 5, 15811. doi:10.1038/srep15811.

Goryluk, A., H. Rekosz-Burlaga, and M. Blaszczyk. 2009. Isolation and characterization of bacterial endophytes of *Chelidonium majus* L. Polish J Microbiol. 58, 355-361.

Gupta, G., Panwar, J., Akhtar, M. S., and Jha, P. N. (2012). “Endophytic nitrogen-fixing bacteria as biofertilizer,” in Sustainable Agriculture Reviews, E. Lichtfouse (ed.), pp. 183–221. Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-5449-2_8.

Gyaneshwar, P., James, E.K., Mathan, N., Reddy, P.M., Reinhold-Hurek, B., Ladha, J.K. (2001). Endophytic colonization of rice by a diazotrophic strain of *Serratia marcescens*. J Bacteriol 183, 2634-2645.

Hall, T.A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT", in: Nucleic acids symposium series, 95-98.

Hammer, Ø., Harper, D., and Ryan, P. (2001). Past: paleontological statistics software package for education and data analysis. Paleontología Electrónica 4, 1-9. URL:< http://palaeo-electronica org/2001_1/past/issue1_01 html.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol. 166, 557–580.

Hardoim, P. R., Andreote, F. D., Reinhold-Hurek, B., Sessitsch, A., van Overbeek, L. S., and van Elsas, J. D. (2011). Rice root-associated bacteria: insights into community structures across 10 cultivars. FEMS Microbiol Ecol. 77, 154–164. doi:10.1111/j.1574-6941.2011.01092.x.

Hardoim, P. R., van Overbeek, L. S., Berg, G., Pirttilä, A. M., Compant, S., Campisano, A., et al. (2015). The Hidden World within Plants: Ecological and evolutionary considerations for defining functioning of microbial endophytes. Microbiol Mol Biol Rev. 79, 293–320. doi:10.1128/MMBR.00050-14.

Hardoim, P.R., van Overbeek, L.S., and van Elsas, J.D. (2008). Properties of bacterial endophytes and their proposed role in plant growth. Trends Microbiol. 16, 463-471.

Hartmann, A., Rothballer, M., Hense, B. A., and Schräder, P. (2014). Bacterial quorum sensing compounds are important modulators of microbe-plant interactions. Front Plant Sci. 5. doi:10.3389/fpls.2014.00131.

Hoagland, D.R., and Arnon, D.I. (1950). The water-culture method for growing plants without soil. Circular. California Agricultural Experiment Station. 347, 1-32.

Hocher, V., Alloisio, N., Bogusz, D., Normand, P. (2011). Early signaling in actinorhizal symbioses. Plant Signal Behav. 6: 1377-1379.

Huang, J-S. (1986). Ultrastructure of bacterial penetration in plants. Annu Rev Phytopathol. 24, 141-157.

Hurek, T., and Reinhold-Hurek, B. (2003). *Azoarcus* sp. strain BH72 as a model for nitrogen-fixing grass endophytes. J Biotechnol. 106, 169–178.

Husemann, P., and Stoye, J. (2010). r2cat: synteny plots and comparative assembly. Bioinformatics. 26, 570-571. doi: 10.1093/bioinformatics/btp690.

Inoue, A., Murata, Y., Takahashi, H., Tsuji, N., Fujisaki, S., and Kato, J.-I. (2008). Involvement of an essential gene, *mviN*, in murein synthesis in *Escherichia coli*. J Bacteriol. 190, 7298-7301.

Ionescu, M., Franchini, A., Egli, T., and Belkin, S. (2008). Induction of the *yjbEFGH* operon is regulated by growth rate and oxygen concentration. Arch Microbiol. 189, 219-226. doi: 10.1007/s00203-007-0311-0.

Istivan, T.S., and Coloe, P.J. (2006). Phospholipase A in Gram-negative bacteria and its role in pathogenesis. Microbiology. 152, 1263-1274. doi: 10.1099/mic.0.28609-0.

Iuchi, S., Matsuda, Z., Fujiwara, T., and Lin, E.C. (1990). The *arcB* gene of *Escherichia coli* encodes a sensor-regulator protein for anaerobic repression of the *arc* modulon. Mol Microbiol. 4, 715-727.

James, E., Olivares, F., Baldani, J., and Döbereiner, J. (1997). *Herbaspirillum*, an endophytic diazotroph colonizing vascular tissue 3*Sorghum bicolor* L. Moench. J Exp Bot. 48, 785-798.

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. Nature. 444, 323–329. doi:10.1038/nature05286.

Jousset, A., Rochat, L., Lanoue, A., Bonkowski, M., Keel, C., Scheu, S. (2010). Plants respond to pathogen infection by enhancing the antifungal gene expression of root-associated bacteria. Mol Plant Microbe Interact. 24, 352-358. doi: 10.1094/MPMI-09-10-0208.

Junker, C., Draeger, S., and Schulz, B. (2012). A fine line – endophytes or pathogens in *Arabidopsis thaliana*. Fungal Ecol. 5, 657–662. doi:10.1016/j.funeco.2012.05.002.

Kanamaru, K., Kanamaru, K., Tatsuno, I., Tobe, T., and Sasakawa, C. (2000). *SdiA*, an *Escherichia coli* homologue of quorum-sensing regulators, controls the expression of virulence factors in enterohaemorrhagic *Escherichia coli* O157:H7. Mol Microbiol. 38, 805-816.

Kaul, S., Sharma, T., and K. Dhar, M. (2016). “Omics” Tools for better understanding the plant–endophyte interactions. Front Plant Sci. 7, 955. doi:10.3389/fpls.2016.00955.

Khan, A., Hamayun, M., Kang, S.-M., Kim, Y.-H., Jung, H.-Y., Lee, J.-H., et al. (2012). Endophytic fungal association via gibberellins and indole acetic acid can improve plant growth under abiotic stress: an example of *Paecilomyces formosus* LHL10. BMC Microbiol. 12, 3. doi:10.1186/1471-2180-12-3.

Kinkel, L. L., Bakker, M. G., and Schlatter, D. C. (2011). A coevolutionary framework for managing disease-suppressive soils. Annu Rev Phytopathol. 49, 47–67. doi:10.1146/annurev-phyto-072910-095232.

Kogel, K.H., Franken, P., Huckelhoven, R. (2006). Endophyte or parasite--what decides? Curr Opin Plant Biol. 9, 358-363. doi: 10.1016/j.pbi.2006.05.001.

Kube, M., Migdoll, A.M., Gehring, I., Heitmann, K., Mayer, Y., Kuhl, H., et al. (2010). Genome comparison of the epiphytic bacteria *Erwinia billingiae* and *E. tasmaniensis* with the pear pathogen *E. pyrifoliae*. BMC Genomics. 11, 393. doi: 10.1186/1471-2164-11-393.

Li, L., Stoeckert, C.J., Jr., and Roos, D.S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13, 2178-2189. doi: 10.1101/gr.1224503.

Lindquist, S., Weston-Hafer, K., Schmidt, H., Pul, C., Korfmann, G., Erickson, J., et al. (1993). *AmpG*, a signal transducer in chromosomal beta-lactamase induction. Mol Microbiol. 9, 703-715.

Linhartová, I., Bumba, L., Mašín, J., Basler, M., Osička, R., Kamanová, J., et al. (2010). RTX proteins: a highly diverse family secreted by a common mechanism. FEMS Microbiol Rev. 34, 1076-1112.

Lomovskaya, O., and Lewis, K. (1992). *emr*, an *Escherichia coli* locus for multidrug resistance. Proc Natl Acad Sci U.S.A. 89, 8938-8942.

Long, S.R. (1996). *Rhizobium symbiosis*: nod factors in perspective. Plant Cell. 8, 1885-1898. doi: 10.1105/tpc.8.10.1885.

Loper, J. E., Hassan, K. A., Mavrodi, D. V., Davis, E. W., Lim, C. K., Shaffer, B. T., et al. (2012). Comparative genomics of plant-associated *Pseudomonas* spp.: insights into diversity and inheritance of traits involved in multitrophic interactions. PLoS Genet. 8, e1002784. doi:10.1371/journal.pgen.1002784.

Lòpez-Fernàndez, S., Sonego, P., Moretto, M., Pancher, M., Engelen, K., Pertot, I., et al. (2015). Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. Front. Microbiol. 6., 419 doi:10.3389/fmicb.2015.00419.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., and Liu, Y. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. Gigascience 1, 18. doi: 10.1186/2047-217X-1-18.

Maddocks, S.E., and Oyston, P.C. (2008). Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology. 154, 3609-3623. doi: 10.1099/mic.0.2008/022772-0.

Malacarne, G., Vrhovsek, U., Zulini, L., Cestaro, A., Stefanini, M., Mattivi, F., Delledonne, M., Velasco, R., Moser, C. (2011). Resistance to *Plasmopara viticola* in a grapevine segregating population is associated with stilbenoid accumulation and with specific host transcriptional responses. BMC Plant Biol. 11, 114. doi: 10.1186/1471-2229-11-114.

Malinowski D, Alloush G, Belesky D (1998) Evidence for chemical changes on the root surface of tall fescue in response to infection with the fungal endophyte *Neotyphodium coenophialum*. Plant Soil. 205, 1-12. doi: 10.1023/A:1004331932018.

Manz, W., Amann, R., Ludwig, W., Wagner, M., Schleifer, K-H. (1992) Phylogenetic oligodeoxynucleotide probes for the major subclasses of proteobacteria: problems and solutions. Syst Appl Microbiology. 15, 593-600.

Marasco, R., Rolli, E., Fusi, M., Cherif, A., Abou-Hadid, A., El-Bahairy, U., et al. (2013). Plant growth promotion potential is equally represented in diverse grapevine root-associated bacterial communities from different biopedoclimatic environments. BioMed Res. Int. 2013, 1–17. doi:10.1155/2013/491091.

Matsumura, E. E., Secco, V. A., Moreira, R. S., dos Santos, O. J. A. P., Hungria, M., and de Oliveira, A. L. M. (2015). Composition and activity of endophytic bacterial communities in field-grown maize plants inoculated with *Azospirillum brasilense*. Ann Microbiol. 65, 2187–2200. doi:10.1007/s13213-015-1059-4.

Matthysse, A.G., White, S., and Lightfoot, R. (1995). Genes required for cellulose synthesis in *Agrobacterium tumefaciens*. J Bacteriol. 177, 1069-1075.

Mccandlish, A.C., and Silhavy, T.J. (2007). Sugar-coating bacteria with lipopolysaccharides. Microbe. 2, 289.

McCullen, C.A., Binns, A.N. (2006). *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. Annu Rev Cell Dev Biol. 22, 101-127.

McCutcheon, J.P., and Moran, N.A. (2012). Extreme genome reduction in symbiotic bacteria. Nat Rev Microbiol. 10, 13-26. doi: 10.1038/nrmicro2670.

Mercado-Blanco, J., and Lugtenberg, B. (2014). Biotechnological applications of bacterial endophytes. Curr Biotechnol. 3, 60–75. doi:10.2174/22115501113026660038.

Meysman, P., Sanchez-Rodriguez, A., Fu, Q., Marchal, K., and Engelen, K. (2013). Expression divergence between *Escherichia coli* and *Salmonella enterica* serovar Typhimurium reflects their lifestyles. Mol Biol Evol. 30, 1302-1314. doi: 10.1093/molbev/mst029.

Mitter, B., Petric, A., Shin, M. W., Chain, P. S. G., Hauberg-Lotte, L., Reinhold-Hurek, B., et al. (2013). Comparative genome analysis of *Burkholderia phytofirmans* PsJN reveals a wide spectrum of endophytic lifestyles based on interaction strategies with host plants. Front Plant Sci 4, 120. doi:10.3389/fpls.2013.00120.

Mo, Y. Y., and Gross, D. C. (1991). Plant signal molecules activate the *syrB* gene, which is required for syringomycin production by *Pseudomonas syringae* pv. syringae. J Bacteriol. 173, 5784–5792. doi:10.1128/jb.173.18.5784-5792.1991.

Møller, T. S. B., Overgaard, M., Nielsen, S. S., Bortolaia, V., Sommer, M. O. A., Guardabassi, L., et al. (2016). Relation between *tetR* and *tetA* expression in tetracycline resistant *Escherichia coli*. BMC Microbiol. 16. doi:10.1186/s12866-016-0649-z.

Moore, L.W., Chilton, W.S., and Canfield, M.L. (1997). Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. Appl Environ Microbiol. 63, 201-207.

Mueller, U. G., and Sachs, J. L. (2015). Engineering microbiomes to improve plant and animal health. Trends Microbiol. 23, 606–617. doi:10.1016/j.tim.2015.07.009.

Munkelt, D., Grass, G., and Nies, D.H. (2004). The chromosomally encoded cation diffusion facilitator proteins DmeF and FieF from *Wautersia metallidurans* CH34 are transporters of broad metal specificity. J Bacteriol. 186, 8036-8043.

Musetti, R., A. Vecchione, L. Stringher, S. Borselli, L. Zulini, C. Marzani, et al. (2006). Inhibition of sporulation and ultrastructural alterations of grapevine downy mildew by the endophytic fungus *Alternaria alternata*. Phytopathology. 96, 689-698.

Oldroyd, G. E. D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat Rev Microbiol. 11, 252–263. doi:10.1038/nrmicro2990.

Oteino, N., Lally, R. D., Kiwanuka, S., Lloyd, A., Ryan, D., Germaine, K. J., et al. (2015). Plant growth promotion induced by phosphate solubilizing endophytic *Pseudomonas isolates*. Front Microbiol. 6. doi:10.3389/fmicb.2015.00745.

Ozer, E.A., Allen, J.P., and Hauser, A.R. (2014). Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt. BMC Genomics. 15, 737. doi: 10.1186/1471-2164-15-737.

Pancher, M., Ceol, M., Corneo, P. E., Longa, C. M. O., Yousaf, S., Pertot, I., et al. (2012). Fungal endophytic communities in grapevines (*Vitis vinifera* L.) respond to crop management. Appl. Environ Microbiol. 78, 4308–4317. doi:10.1128/AEM.07655-11.

Pandey, D.P., and Gerdes, K. (2005). Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res. 33, 966-976. doi: 10.1093/nar/gki201.

Park, D.H., Thapa, S.P., Choi, B.S., Kim, W.S., Hur, J.H., Cho, J.M., et al. (2011). Complete genome sequence of Japanese *Erwinia* strain ejp617, a bacterial shoot blight pathogen of pear. J Bacteriol. 193, 586-587. doi: 10.1128/jb.01246-10.

Parsons, J.F., Song, F., Parsons, L., Calabrese, K., Eisenstein, E., and Ladner, J.E. (2004). Structure and function of the phenazine biosynthesis protein PhzF from *Pseudomonas fluorescens* 2-79. Biochemistry 43, 12427-12435. doi: 10.1021/bi049059z.

Partida-Martínez, L. P., and Heil, M. (2011). The microbe-free plant: fact or artifact? Front Plant Sci. 2. doi:10.3389/fpls.2011.00100.

Pažoutová, S., Follert, S., Bitzer, J., Keck, M., Surup, F., Srutka, P. et al. (2013). A new endophytic insect-associated *Daldinia* species, recognised from a comparison of secondary metabolite profiles and molecular phylogeny. Fung Div. 60, 107. doi:10.1007/s13225-013-0238-5

Perazzolli, M., Antonielli, L., Storari, M., Puopolo, G., Pancher, M., Giovannini, O., et al. (2014). Resilience of the natural phyllosphere microbiota of the grapevine to chemical and biological pesticides. Appl Environ Microbiol. 80, 3585–3596. doi:10.1128/AEM.00415-14.

Perrine-Walker, F., Doumas, P., Lucas, M., Vaissayre, V., Beauchemin, N.J., Band, L.R., et al. (2010). Auxin carriers localization drives auxin accumulation in plant cells infected by *Frankia* in *Casuarina glauca* actinorhizal nodules. Plant Physiol. 154, 1372-1380. doi: 10.1104/pp.110.163394.

Perrine-Walker, F. M., Prayitno, J., Rolfe, B. G., Weinman, J. J., and Hocart, C. H. (2007). Infection process and the interaction of rice roots with rhizobia. J Exp Bot. 58, 3343–3350. doi:10.1093/jxb/erm181.

Pezet, R., Gindro, K., Viret, O., Richter, H. (2004) Effects of resveratrol, viniferins and pterostilbene on *Plasmopara viticola* zoospore mobility and disease development. Vitis. 43, 145-148.

Pinto, C., and Gomes, A. C. (2016). *Vitis vinifera* microbiome: from basic research to technological development. BioControl. 61, 243–256. doi:10.1007/s10526-016-9725-4.

Pinto, C., Pinho, D., Sousa, S., Pinheiro, M., Egas, C., and C. Gomes, A. (2014). Unravelling the diversity of grapevine microbiome. PLoS ONE. 9, e85622. doi:10.1371/journal.pone.0085622.

Pirttilä, A. M., Joensuu, P., Pospiech, H., Jalonen, J., and Hohtola, A. (2004). Bud endophytes of Scots pine produce adenine derivatives and other compounds that affect morphology and mitigate browning of callus cultures. Physiol Plant. 121, 305–312. doi:10.1111/j.0031-9317.2004.00330.x.

Poole, K., and Braun, V. (1988). Iron regulation of *Serratia marcescens* hemolysin gene expression. Infect Immun. 56, 2967-2971.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. [Online].

Raman, A., Wheatley, W., and Popay, A. (2012). Endophytic fungus-vascular plant-insect interactions. Environ Entomol. 41, 433–447. doi:10.1603/EN11317.

Reineke, A., and Thiéry, D. (2016). Grapevine insect pests and their natural enemies in the age of global warming. J Pest Sci. 89, 313–328. doi:10.1007/s10340-016-0761-8.

Reinhold-Hurek, B., and Hurek, T. (2011). Living inside plants: bacterial endophytes. Curr Opin Plant Biol. 14, 435-443. doi: 10.1016/j.pbi.2011.04.004.

Remus-Emsermann, M.N., Kim, E.B., Marco, M.L., Tecon, R., and Leveau, J.H. (2013). Draft genome sequence of the phyllosphere model bacterium *Pantoea agglomerans* 299R. Genome announc. 1, e00036-00013.

Ren, Y., Ren, Y., Zhou, Z., Guo, X., Li, Y., Feng, L., et al. (2010). Complete genome sequence of Enterobacter cloacae subsp. cloacae type strain ATCC 13047. J. Bacteriol. 192, 2463-2464.

Robinson, R. J., Fraaije, B. A., Clark, I. M., Jackson, R. W., Hirsch, P. R., and Mauchline, T. H. (2016). Endophytic bacterial community composition in wheat (*Triticum aestivum*) is determined by plant tissue type, developmental stage and soil nutrient availability. Plant Soil. 405, 381–396. doi:10.1007/s11104-015-2495-4.

Rodriguez, R., and Redman, R. (2008). More than 400 million years of evolution and some plants still can't make it on their own: plant stress tolerance via fungal symbiosis. J Exp Bot. 59, 1109–1114. doi:10.1093/jxb/erm342.

Rodríguez-Navarro, D. N., Dardanelli, M. S., and Ruíz-Saenz, J. E. (2007). Attachment of bacteria to the roots of higher plants. FEMS Microbiol. Lett. 272, 127–136. doi:10.1111/j.1574-6968.2007.00761.x.

Ryan, R. P., Germaine, K., Franks, A., Ryan, D. J., and Dowling, D. N. (2008). Bacterial endophytes: recent developments and applications. FEMS Microbiol. Lett. 278, 1–9. doi:10.1111/j.1574-6968.2007.00918.x.

Sagaram, U. S., DeAngelis, K. M., Trivedi, P., Andersen, G. L., Lu, S.-E., and Wang, N. (2009). Bacterial diversity analysis of huanglongbing pathogen-infected citrus, using PhyloChip arrays and 16S rRNA gene clone library sequencing. Appl Environ Microbiol. 75, 1566–1574. doi:10.1128/AEM.02404-08.

Saikkonen, K. (2004). Evolution of endophyte-plant symbioses. Trends Plant Sci. 9, 275–280. doi:10.1016/j.tplants.2004.04.005.

Sakai, D., and Komano, T. (2000). The *pilL* and *pilN* genes of IncI1 plasmids R64 and ColIb-P9 encode outer membrane lipoproteins responsible for thin pilus biogenesis. Plasmid. 43, 149-152. doi: 10.1006/plas.1999.1434.

Samad, A., Trognitz, F., Compant, S., Antonielli, L., and Sessitsch, A. (2017). Shared and host-specific microbiome diversity and functioning of grapevine and accompanying weed plants: Microbial communities associated with grapevine and vineyard weeds. Environ Microbiol. doi:10.1111/1462-2920.13618.

Sana, T.G., Soscia, C., Tonglet, C.M., Garvis, S., and Bleves, S. (2013). Divergent control of two type VI secretion systems by RpoN in *Pseudomonas aeruginosa*. PLoS ONE. 8, e76030. doi: 10.1371/journal.pone.0076030.

Santoyo, G., Moreno-Hagelsieb, G., del Carmen Orozco-Mosqueda, M., and Glick, B. R. (2016). Plant growth-promoting bacterial endophytes. Microbiol Res. 183, 92–99. doi:10.1016/j.micres.2015.11.008.

Schaefer, A. L., Lappala, C. R., Morlen, R. P., Pelletier, D. A., Lu, T.-Y. S., Lankford, P. K., et al. (2013). LuxR- and LuxI-Type quorum-sensing circuits are prevalent in members of the *Populus deltoides* microbiome. Appl Environ Microbiol. 79, 5745–5752. doi:10.1128/AEM.01417-13.

Scherling C, Ulrich K, Ewald D, Weckwerth W (2009) A metabolic signature of the beneficial interaction of the endophyte *Paenibacillus* sp. isolate and *in vitro*-grown poplar plants revealed by metabolomics. Mol Plant Microbe Interact. 22, 1032-1037. doi: 10.1094/mpmi-22-8-1032.

Schlieker, C., Zentgraf, H., Dersch, P., and Mogk, A. (2005). ClpV, a unique Hsp100/Clp member of pathogenic proteobacteria. Biol Chem. 386, 1115-1127. doi: 10.1515/bc.2005.128.

Schmidt, H., and Eickhorst, T. (2014). Detection and quantification of native microbial populations on soil-grown rice roots by catalyzed reporter deposition-fluorescence *in situ* hybridization. FEMS Microbiol Ecol. 87, 390–402. doi:10.1111/1574-6941.12232.

Schmidt, M., Balsanelli, E., Faoro, H., Cruz, L. M., Wassem, R., de Baura, V. A., et al. (2012). The type III secretion system is necessary for the development of a pathogenic and endophytic interaction between *Herbaspirillum rubrisubalbicans* and Poaceae. BMC Microbiol. 12, 98. doi:10.1186/1471-2180-12-98.

Schulz, B., and Boyle, C. (2005). The endophytic continuum. Mycol. Res. 109, 661-686.

Schulz, B., and Boyle, C. (2006). "What are Endophytes?" in Microbial Root Endophytes, B. J. E. Schulz, C. J. C. Boyle, and T. N. Sieber (eds), pp. 1–13. Berlin, Heidelberg: Springer Berlin Heidelberg Available at: http://link.springer.com/10.1007/3-540-33526-9_1 [Accessed January 17, 2017].

Schulz, B., Haas, S., Junker, C., Andrée, N., and Schobert, M. (2015). Fungal endophytes are involved in multiple balanced antagonisms. Curr Sci. 109, 39-45

Schulz, B., Römmert, A.-K., Dammann, U., Aust, H-J., and Strack, D. (1999). The endophyte-host interaction: a balanced antagonism? Mycol Res. 103, 1275-1283.

Sebaihia, M., Bocsanczy, A., Biehl, B., Quail, M., Perna, N., Glasner, J., et al. (2010). Complete genome sequence of the plant pathogen *Erwinia amylovora* strain ATCC 49946. J Bacteriol. 192, 2020-2021.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics. 30, 2068-2069. doi: 10.1093/bioinformatics/btu153.

Sessitsch, A., Hardoim, P., Döring, J., Weilharter, A., Krause, A., Woyke, T., et al. (2012). Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. Mol Plant Microbe Interact. 25, 28–36. doi:10.1094/MPMI-08-11-0204.

Shepard, W., Soutourina, O., Courtois, E., England, P., Haouz, A., and Martin-Verstraete, I. (2011). Insights into the Rrf2 repressor family–the structure of CymR, the global cysteine regulator of *Bacillus subtilis*. FEBS J. 278, 2689-2701.

Sheppard, S. K., Colles, F. M., McCarthy, N. D., Strachan, N. J. C., Ogden, I. D., Forbes, K. J., et al. (2011). Niche segregation and genetic structure of *Campylobacter jejuni* populations from wild and agricultural host species. Mol Ecol. 20, 3484–3490. doi:10.1111/j.1365-294X.2011.05179.x.

Shiba, Y., Matsumoto, K., and Hara, H. (2006). DjlA negatively regulates the Rcs signal transduction system in *Escherichia coli*. Genes Genet Syst. 81, 51-56. doi: 10.1266/ggs.81.51.

Shidore T, Dinse T, Ohrlein J, Becker A, Reinhold-Hurek B (2012) Transcriptomic analysis of responses to exudates reveal genes required for rhizosphere competence of the endophyte *Azoarcus* sp. strain BH72. Environ Microbiol. 14, 2775-2787. doi: 10.1111/j.1462-2920.2012.02777.x.

Shimada, T., Fujita, N., Yamamoto, K., and Ishihama, A. (2011). Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. PLoS ONE. 6, e20081. doi: 10.1371/journal.pone.0020081.

Shin, S.H., Kim, S., Kim, J.Y., Lee, S., Um, Y., Oh, M.K., et al. (2012). Complete genome sequence of *Enterobacter aerogenes* KCTC 2190. J Bacteriol. 194, 2373-2374. doi: 10.1128/jb.00028-12.

Silby, M. W., Cerdeño-Tárraga, A. M., Vernikos, G. S., Giddens, S. R., Jackson, R. W., Preston, G. M., et al. (2009). Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. Genome Biol. 10, R51. doi:10.1186/gb-2009-10-5-r51.

Smits, T.H., Rezzonico, F., Kamber, T., Goesmann, A., Ishimaru, C.A., Stockwell, V.O., et al. (2010). Genome sequence of the biocontrol agent *Pantoea vagans* strain C9-1. J. Bacteriol. 192, 6486-6487.

Sorensen, K.I., and Hove-Jensen, B. (1996). Ribose catabolism of *Escherichia coli*: characterization of the *rpiB* gene encoding ribose phosphate isomerase B and of the *rpiR* gene, which is involved in regulation of *rpiB* expression. J Bacteriol. 178, 1003-1011.

Spence, C., and Bais, H. (2013). "Probiotics for plants: rhizospheric microbiome and plant fitness," in Molecular Microbial Ecology of the Rhizosphere, F. J. de Bruijn (ed.), pp 713–721. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi:10.1002/9781118297674.ch67.

Sturz, A.V., Christie, B.R., Nowak, J. (2000). Bacterial endophytes: potential role in developing sustainable systems of crop production. Crr cr Rev Plant Sci 19, 1-30. doi: 10.1080/07352680091139169.

Suen, G., Goldman, B. S., and Welch, R. D. (2007). Predicting prokaryotic ecological niches using genome sequence analysis. PLoS ONE. 2, e743. doi:10.1371/journal.pone.0000743.

Sugawara, M., Epstein, B., Badgley, B. D., Unno, T., Xu, L., Reese, J., et al. (2013). Comparative genomics of the core and accessory genomes of 48 *Sinorhizobium* strains comprising five genospecies. Genome Biol. 14, R17. doi:10.1186/gb-2013-14-2-r17.

Taghavi, S., Garafola, C., Monchy, S., Newman, L., Hoffman, A., Weyens, N., et al. (2009). Genome survey and characterization of endophytic bacteria exhibiting a beneficial effect on growth and development of Poplar trees. Appl Environ Microbiol. 75, 748–757. doi:10.1128/AEM.02239-08.

Taghavi, S., van der Lelie, D., Hoffman, A., Zhang, Y.-B., Walla, M. D., Vangron sveld, J., et al. (2010). Genome sequence of the plant growth promoting endophytic bacterium *Enterobacter* sp. 638. PLoS Genet. 6, e1000943. doi:10.1371/journal.pgen.1000943.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.

Tanaka, A., Christensen, M.J., Takemoto, D., Park, P., Scott, B. (2006). Reactive oxygen species play a role in regulating a fungus-perennial ryegrass mutualistic interaction. Plant cell. 18, 1052-1066. doi: 10.1105/tpc.105.039263.

Tian, B.-Y., Cao, Y., and Zhang, K.-Q. (2015). Metagenomic insights into communities, functions of endophytes, and their associates with infection by root-knot nematode, Meloidogyne incognita, in tomato roots. Sci. Rep. 5, 17087. doi:10.1038/srep17087.

Tian, C. F., Zhou, Y. J., Zhang, Y. M., Li, Q. Q., Zhang, Y. Z., Li, D. F., et al. (2012). Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. Proc Natl Acad Sci. 109, 8629–8634. doi:10.1073/pnas.1120436109.

Tintjer, T., Leuchtmann, A., and Clay, K. (2008). Variation in horizontal and vertical transmission of the endophyte *Epichloë elymi* infecting the grass *Elymus hystrix*. New Phytol. 179, 236–246. doi:10.1111/j.1469-8137.2008.02441.x.

Torres, M. S., White, J. F., Zhang, X., Hinton, D. M., and Bacon, C. W. (2012). Endophyte-mediated adjustments in host morphology and physiology and effects on host fitness traits in grasses. Fungal Ecol. 5, 322–330. doi:10.1016/j.funeco.2011.05.006.

Trdá, L., Boutrot, F., Claverie, J., Brulè, D., Dorey, S., and Poinssot, B. (2015). Perception of pathogenic or beneficial bacteria and their evasion of host immunity: pattern recognition receptors in the frontline. Front. Plant Sci. 6. doi:10.3389/fpls.2015.00219.

Trdá, L., Fernandez, O., Boutrot, F., Héloir, M.-C., Kelloniemi, J., Daire, X., et al. (2014). The grapevine flagellin receptor VvFLS2 differentially recognizes flagellin-derived epitopes from the endophytic growth-promoting bacterium *Burkholderia phytofirmans* and plant pathogenic bacteria. New Phytol. 201, 1371–1384. doi:10.1111/nph.12592.

Tritt, A., Eisen, J.A., Facciotti, M.T., and Darling, A.E. (2012). An Integrated pipeline for *de novo* assembly of microbial genomes. PLoS ONE. 7, e42304. doi: 10.1371/journal.pone.0042304.

Truyens, S., Weyens, N., Cuypers, A., and Vangronsveld, J. (2015). Bacterial seed endophytes: genera, vertical transmission and interaction with plants: bacterial seed endophytes. Environ Microbiol Rep. 7, 40–50. doi:10.1111/1758-2229.12181.

Turner, T. R., James, E. K., and Poole, P. S. (2013). The plant microbiome. Genome Biol. 14. doi:10.1186/gb-2013-14-6-209.

Vrhovsek, U., Masuero, D., Gasperotti, M., Franceschi, P., Caputi, L., Viola, R., Mattivi, F. (2012). A versatile targeted metabolomics method for the rapid quantification of multiple classes of phenolics in fruits and beverages. J Agr Food Chem. 60, 8831-8840. doi: 10.1021/jf2051569.

Wall, L., Christiansen, T., and Orwant, J. (2000). Programming Perl. California: O'Reilly & Associates, Inc.

Wallner, G., Amann, R., and Beisker, W. (1993). Optimizing fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry. 14, 136-143. doi: 10.1002/cyto.990140205.

West, E. R., Cother, E. J., Steel, C. C., and Ash, G. J. (2010). The characterization and diversity of bacterial endophytes of grapevine. Can J Microbiol. 56, 209–216. doi:10.1139/W10-004.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., et al. (2007). Database resources of the national center for biotechnology information. Nucleic Acids res. 35, D5-D12.

White, J.F., Jr., Torres, M.S. (2010). Is plant endophyte-mediated defensive mutualism the result of oxidative stress protection? Physiol plantarum. 138, 440-446. doi: 10.1111/j.1399-3054.2009.01332.x.

White, T.J., T. Bruns, S.W. Lee, and J.W. Taylor. (1990). “Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics”. In PCR Protocols: A guide to methods and applications. Innis, M.A., Gelfand, D.H., Sninsky, J.J., and White, T.J. (eds.), pp. 315-322. New York: Academic Press Inc.

Whiteside, S. A., Razvi, H., Dave, S., Reid, G., and Burton, J. P. (2015). The microbiome of the urinary tract—a role beyond infection. Nat Rev Urol. 12, 81–90. doi:10.1038/nrurol.2014.361.

Winans, S.C. (1992) Two-way chemical signaling in *Agrobacterium*-plant interactions. Microbiol Rev. 56, 12-31.

Wu, C. H., Bernard, S. M., Andersen, G. L., and Chen, W. (2009). Developing microbe-plant interactions for applications in plant-growth promotion and disease control, production of useful compounds, remediation and carbon sequestration. Microb Biotechnol. 2, 428–440. doi:10.1111/j.1751-7915.2009.00109.x.

Wuana, R. A., and Okieimen, F. E. (2011). Heavy metals in contaminated soils: a review of sources, chemistry, risks and best available strategies for remediation. ISRN Ecol. 1–20. doi:10.5402/2011/402647.

Xu, X-H., Su, Z-Z., Wang, C., Kubicek, C.P., Feng, X.-X., Mao, L.-J., et al. (2014). The rice endophyte *Harpophora oryzae* genome reveals evolution from a pathogen to a mutualistic endophyte. Sci Rep. 4. doi: 10.1038/srep05783

Yamaguchi, Y., Park, J.-H., and Inouye, M. (2011). Toxin-antitoxin systems in bacteria and archaea. Annu Rev Genet. 45, 61–79. doi:10.1146/annurev-genet-110410-132412.

Yang, L., Yinhu, L., Su, Y., Hui, W., Yanhua, C., Jie, L., et al. (2015). Diversity and distribution of endophytic bacterial community in the Noni (Morinda citrifolia L.) plant. Afr J Microbiol Res. 9, 1649–1657. doi:10.5897/AJMR2015.7443.

Yeh, K.-C., Peck, M. C., and Long, S. R. (2002). Luteolin and GroESL modulate *in vitro* activity of NodD. J Bacteriol. 184, 525–530. doi:10.1128/JB.184.2.525-530.2002.

Yousaf, S., Bulgari, D., Bergna, A., Pancher, M., Quaglino, F., Casati, P., et al. (2014). Pyrosequencing detects human and animal pathogenic taxa in the grapevine endosphere. Front Microbiol. 5. doi:10.3389/fmicb.2014.00327.

Yu, L., Sang, W., Wu, M.-D., Zhang, J., Yang, L., Zhou, Y.-J., et al. (2015). Novel hypovirulence-associated RNA Mycovirus in the plant-pathogenic fungus *Botrytis cinerea*: molecular and biological characterization. Appl Environ Microbiol. 81, 2299–2310. doi:10.1128/AEM.03992-14.

Zarraonaindia, I., and Gilbert, J. (2015). Understanding grapevine-microbiome interactions: implications for viticulture industry. Microb Cell. 2, 171–173. doi:10.15698/mic2015.05.204.

Zeller, S. L., Kalinina, O., and Schmid, B. (2013). Costs of resistance to fungal pathogens in genetically modified wheat. J Plant Ecol. 6, 92–100. doi:10.1093/jpe/rts013.

Zhang, J., Subramanian, S., Stacey, G., and Yu, O. (2009). Flavones and flavonols play distinct critical roles during nodulation of *Medicago truncatula* by *Sinorhizobium meliloti*. Plant J. 57, 171–183. doi:10.1111/j.1365-313X.2008.03676.x.

Zhou, C.E., Smith, J., Lam, M., Zemla, A., Dyer, M.D., and Slezak, T. (2007). MvirDB--a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. Nucleic Acids Res. 35, D391-394. doi: 10.1093/nar/gkl791.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., and Wishart, D.S. (2011). PHAST: A Fast Phage Search Tool. Nucleic Acids Res. doi: 10.1093/nar/gkr485.

Zook, D. „Symbiosis - evolution´s Co-author“. In Reticulate evolution. Gontier, N. (ed.), pp. 41-80. Springer International publisher Switzerland. Available at http://www.springer.com/gp/book/9783319163444. [Accessed March 18, 2017]

Zinniel, D.K., Lambrecht, P., Harris, N.B., Feng , Z., Kuczmarski , D., Higley et al. 2002. Isolation and characterization of endophytic colonizing bacteria from agronomic crops and prairie plants. Appl Environ Microbiol. 68, 2198-2208.

CURRICULUM VITAE

Personal data	
Name:	Juan Sebastian Lopez Fernandez
Birth date and place:	28.07.1986 – Bogota, Colombia
Family status:	Single

EDUCATION

1997-2003	Walt Whitman School. Bogotá, Colombia
2004-2010	National University of Colombia Bachelor in Biology. Major in Bacteriology Studies of the bacterial diversity in the coral *Pseudopterogorgia elisabetae*
2010 – 2012	National University of Colombia *Magister Scientiae*. Microbiology Evaluation of cell-to-cell signaling between *Pasteurella multocida* and *Bordetella bronchiseptica* in an *in vitro* model of rabbit nasal epithelium Supervised by Prof. Dr. Catalina Arévalo *Magna cum laude*
2012 – 2016	PhD student at Fondazione Edmund Mach (Italy) and Technische Universität Braunschweig Supervised by Dr. Andrea Campisano and PD. Dr. Barbara J. Schulz
2016 – Today	Guest student Helmholtz Zentrum für Infektionsforschung Microbial Drugs Department Supervised by Prof. Dr. Marc Stadler and Dr. Frank Surup

OTHER ACTIVITIES

2009 – 2012	Participant in the research group “Comunication and Bacterial Comunities” in Bogota, Colombia
2010 – 2012	Participant in the research group “Marine natural products” in Bogota Colombia
2010 - 2012	Participant in the research group “Animal pathobiology” in Bogota Colombia
2009 – 2012	Teaching assistantships at the Biology department, National University of Colombia

CONFERENCES

September 23rd - 27th. 2010	**XX Latin-American congress of Microbiology and IX meeting of microbiologists** Poster: “Determination of the histological damage produced by the interaction between *Pasteurella multocida* and *Bordetella bronchiseptica* due to extracellular signaling during pathogenesis”

May 23rd – 26th. 2011	**Practical and Theoretical course: "*Quorum Sensing* in plant-associated bacteria"** Conference: "Do *Pasteurella multocida* and *Bordetella bronchiseptica* talk to each other?"
June 20th – 23rd. 2012	**62nd annual conference of the Canadian Society for microbiology.** Poster: "Do *Pasteurella multocida* and *Bordetella bronchiseptica* talk to each other to develop virulence strategies*?"*
November 3d – 2014	**COST action meeting: Risk assessment of endophytes.** Conference: "A whole-genome comparison of virulence traits in endophytic genomes of Enterobacteria".
June 21st -25th. 2015	**International congress "Rhizo4: stretching the interface of life"** Conference: "Endophytic colonization of grapevine by bacteria reveals a metabolic signature suggesting activation of pathways for symbiosis and defense" Poster: "Insect vectors efficiently convey complex endophytic communities across grapevine plants"
November 23rd – 25th. 2015.	**International symposium "miCROPe" Microbe-assisted crop production- opportunities, challenges and needs.** Conference: "Can insect pests be vectors of beneficial endophytes?"
February 10th – 12th. 2016	**COST action FA1405: "using three-way interactions between plants, microbes and arthropods to enhance crop protection and production"** Conference: "Insect mediated transfer of microbial communities across plants
July 17th – 21st. 2016	**XVII congress of the international society for molecular plant-microbe interactions** Conference and poster:"Adaptation of transmissible bacterial communities to multiple hosts: how the sap-feeding insect S*caphoideus titanus* ships bacterial symbionts across grapevine plants" Poster: **"**Fungal Endophytes From Grapevine Have Host-Dependent Levels Of Virulence And Produce Antibiotic Compounds In Dual Cultures
September 19th – 22nd. 2016	**XXII National Congress of the Italian Society of plant pathology** Poster: **"**Comparision Of Three Artificial Methods For The Re-Inoculation Of Bacterial Endophytes In Micropropagated *Malus Domestica* (Borkh) Plantlets"

PUBLICATIONS

Lòpez-Fernàndez, S., Sonego, P., Moretto, M., Pancher, M., Engelen, K., Pertot, I., Campisano, A. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. Frontiers in Microbiology. 2015: 6.

Lòpez-Fernàndez, S., Compant, S. Vrhovsek, U. Bianchiedi, P L, Sessitsch, A. Pertot, I. Campisano, A. Grapevine colonization by endophytic bacteria shifts secondary metabolism and suggests activation of defense pathways. Plant and Soil. 2015: pp 1-21

Lopez-Fernàndez, S., Mazzoni, V., Pedrazzoli, F., Pertot, I., Campisano, A. A phloem-feeding insect transfers bacterial endophytic communities between grapevine plants. Submitted to Frontiers in Microbiology . January 2017

Campisano, A., Pancher, M., Puopolo, G., Puddu, A., **Lòpez-Fernàndez**, S., Biagini, B., et al. Diversity in endophytic populations reveals functional and taxonomic diversity between wild and domesticated grapevines. American Journal of Enology and Viticulture, ajev-2014.